DOCUMENT RESUME

ED 367 161 FL 021 883

AUTHOR Sadtono, Eugenius, Ed.
TITLE Language Acquisition and the Second/Foreign Language Classroom. Anthology Series 28.
INSTITUTION Southeast Asian Ministers of Education Organization (Singapore). Regional Language Centre.
REPORT NO RELC-P-393-91
PUB DATE 91
NOTE 225p.; Selected papers from the Southeast Asian Ministers of Education Organization (SEAMEO) Regional Language Centre (Singapore, 1991). For individual papers, see FL 021 884-889, FL 021 891, ED 337 047, and ED 338 037.
PUB TYPE Collected Works - General (020) -- Collected Works - Serials (022)

EDRS PRICE MF01/PC09 Plus Postage.
DESCRIPTORS *Attitude Measures; Classroom Techniques; Foreign Countries; Higher Education; High Schools; Interaction; Journal Writing; Language Attitudes; *Language Research; *Learning Theories; Oral Language; *Personality Traits; Research Methodology; Research Needs; *Second Language Instruction; Second Language Learning; Student Characteristics; Student Motivation; Student Role; Teacher Education; Theory Practice Relationship

ABSTRACT

A selection of papers on second language learning includes: "Second Language Acquisition Research in the Language Classroom" (David Nunan); "A Place for Second Language Acquisition in Teacher Development and in Teacher Education Programmes" (Rod Bolitho); "Dimensions in the Acquisition of Oral Language" (Martin Bygate, Don Porter); "The Learner's Effort in the Language Classroom" (N. S. Prabhu); "Diary Studies of Classroom Language Learning: The Doubting Game and the Believing Game" (Kathleen M. Bailey); "Personality and Second-Language Learning: Theory, Research and Practice" (Roger Griffiths); "The Contribution of SLA Theories and Research to Teaching Language" (Andrew S. Cohen, Diane Larsen-Freeman, Elaine Tarone); "The Matched-Guise Technique for Measuring Attitudes and Their Implications for Language Education: A Critical Assessment" (Stephen J. Gaies, Jacqueline D. Beebe); and "The Interaction Hypothesis: A Critical Evaluation" (Rod Ellis). (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 367 161

LANGUAGE ACQUISITION AND THE SECOND/FOREIGN LANGUAGE CLASSROOM

REGIONAL LANGUAGE CENTRE

SEAMEO

Edited by

EUGENIUS SADTONO

ANTHOLOGY SERIES 28

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Wong Kim Weng

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

FL021993

BEST COPY AVAILABLE

LANGUAGE ACQUISITION AND THE SECOND/FOREIGN LANGUAGE CLASSROOM

LANGUAGE ACQUISITION AND THE SECOND/FOREIGN LANGUAGE CLASSROOM

Edited

by

Eugenius Sadtono

Anthology Series 28
Published by
SEAMEO Regional Language Centre

RELC P393-91

Copyright © 1991 by SEAMEO Regional Language Centre
30 Orange Grove Road
Singapore 1025
ISBN 9971-74-040-0
ISSN 0129-8895

CONTENTS

	Page
Foreword	iii
Introduction	iv
Second Language Acquisition Research in the Language Classroom *David Nunan*	1
A Place for Second Language Acquisition in Teacher Development and in Teacher Education Programmes *Rod Bolitho*	25
Dimensions in the Acquisition of Oral Language *Martin Bygate and Don Porter*	38
The Learner's Effort in the Language Classroom *N S Prabhu*	49
Diary Studies of Classroom Language Learning: The Doubting Game and the Believing Game *Kathleen M Bailey*	60

Page

Personality and Second-Language Learning: Theory, Research and Practice 103
Roger Griffiths

The Contribution of SLA Theories and Research to Teaching Language 136
Andrew S Cohen, Diane Larsen-Freeman, and Elaine Tarone

The Matched-Guise Technique for Measuring Attitudes and Their Implications for Language Education: A Critical Assessment 156
Stephen J Gaies and Jacqueline D Beebe

The Interaction Hypothesis: A Critical Evaluation 179
Rod Ellis

FOREWORD

One of the functions of SEAMEO RELC is to collect and disseminate information related to the teaching, learning and use of languages within the region, and information related to research priorities in these areas. The annual RELC Regional Seminar is held to fulfill this function. The theme of the 1991 Regional Seminar is "Language Acquisition and the Second/Foreign Language Classroom" and the papers presented deal with a variety of topics related to the theme which would be useful for teaching, learning and research. It is hoped therefore that the reader would benefit considerably from reading the papers selected for this anthology.

As language acquisition covers a very broad area, for practical purposes we narrow it down to second/foreign language acquisition in the classroom. After all most of us are teachers of second/foreign language in the classroom. Even within the classroom the problems of language acquisition are not necessarily contained or minimised, in fact new problems arise.

The papers in this anthology reveal some new problems, and we hope that we would be able to handle them using some of the suggestions the writers made. In addition, the reader would be able to learn the current state of the art of SLA research from David Nunan and Andrew Cohen et. al, the latest developments and techniques on SLA research from Stephen Gaies and Jacqueline Beebe, as well as from Kathleen Bailey. Rod Ellis' critical evaluation of the Interaction Hypothesis is challenging indeed. Prabhu's introduction of the typology of learner effort is a seminal contribution to SLA. Roger Griffith's insistence on personality as a strong variable in language acquisition will certainly revive our interest in personality theory. Martin Bygates' and Don Porter's research findings on the relationship between task and language development will definitely stimulate further research on the topic, and Rod Bolitho's appeal to SLA researchers to assist teacher trainers to become principled practitioners is worthy of our attention.

To conclude, let me avail myself of this opportunity to express my deepest gratitude to all the paper presenters for their contribution and all members of the organising committee for their hard work which accounted for the success of the seminar and the publication of this anthology.

Edwin Goh
Director

INTRODUCTION

Second Language Acquisition covers a wide range of topics. In the 1991 RELC Regional Seminar on Language Acquisition and the Second/Foreign Language Classroom, the topics varied from theoretical treatment of SLA, quantitative and qualitative research on SLA, to practical teaching-learning strategies in SLA. Due to financial constraints and much to our regret, we cannot publish all of the selected papers we thought would be interesting and thought provoking.

The topics of the articles in this book also vary. David Nunan's paper on Second Language Acquisition Research in the Classroom takes a critical assessment of the current state of SLA research, particularly research intended to provide directions for classroom instruction. To improve research, he suggests five key points to be incorporated into future research design, they are (1) the implementation of more contextualised research, i.e. more classroom-based; (2) an extension of the theoretical bases of research agenda; (3) an extension of the range of research tools, techniques and methods; (4) a re-evaluation of the distinction between process-oriented and product-oriented research; and (5) a more active role for classroom practitioners in applied research.

Rod Bolitho's paper on A Place for Second Language Acquisition in Teacher Development and in Teacher Education Programmes is for all intents and purposes an appeal to SLA researchers to assist pre-service and in-service teacher trainers to become 'principled practitioners' so they would be able to ask the right questions and to arrive at a better understanding of the reasons and explanations of successful language learning, to resist the temptations of panacea and recipe-type solutions, to lay the groundwork for continuing professional development. He goes on to say that trainers need principled support, not confusing messages from the world of research.

In their paper on "Dimensions in the Acquisition of Oral Language", Martin Bygate and Don Porter show the relationships between task and language development. Their paper is based on the result of a small pilot study to answer the following questions: what is the effect of task familiarity on language production; and more generally, in what ways does oral production on a given task improve? The result of the study suggests that there could be at least three different patterns of oral language development: in fluency, in linguistic complexity or in both.

In his paper entitled "The Learner's Effort in the Language Classroom", Prabhu identifies different forms of effort by language learners in the classroom. In addition, he suggests a typology of learner effort, by first introducing the concepts of imitation and cerebration. By cerebration he means the process by which the learner thinks things out. In other words, if by imitation the learner copies the language knower's (the teacher's) behaviour, by cerebration s/he attempts to become a language knower, i.e. an effort to cope with problems: to make sense, to work things out, and to put solutions to test. He then proceeds with the typology of learner effort which comprises reproduction, simulation, construction and deployment; the first two involving imitation and the rest cerebration. The conclusion of the paper deals with his claims of the usefulness of the typology in language education.

"Diary Studies of Classroom Language Learning: The Doubting Game and The Believing Game" is the title of Kathleen Bailey's paper. As an introduction, she discusses the meaning of diary studies. A diary study is an account of a second language experience as recorded in a first-person journal. The diarist can be either a language teacher or a language learner. The main characteristic of diary studies is that they are introspective and can reveal facets of language learning experience which are normally hidden or largely inaccessible to an external observer. She then goes on reviewing a number of language learning diary studies, particularly diary studies with introspective (first-person) analysis and the ones with non-introspective analysis. Following the review, she explains the meaning of the doubting game and the believing game. The doubting game seeks truth by indirection -- by seeking errors, such as the null hypothesis in experimental research, whereas the believing game by affirmation. She further subjects diary studies to the doubting game and expose their problems regarding subjects, data collection and data analysis. Finally she subjects them to the believing game and reveals their benefits for language teaching, language teachers and language learners.

Roger Griffiths's untimely death in Phuket, Thailand (he drowned), is indeed a big loss to us. He was still young and a very promising scholar, and he always drew a big crowd whenever he presented a paper as his topic was always challenging and scintillating.

His swan song here deals with personality and language learning. He acknowledged that personality is only one of the many variables that play a crucial role in contributing to the individual differences in SLA; but he insisted that personality variable should not be easily dismissed in SLA research. A rebirth in research on personality variable in SLA is now growing.

Andrew Cohen, Diane Larsen-Freeman and Elaine Tarone wrote a paper on the contribution of SLA theories and research to teaching language. They argue that results of SLA studies may have contributed to the understanding of language learning, but classroom instruction may not benefit directly from the insights gained from the investigations. This may be due to the fact that Second-Language-Acquisition research agenda is not necessarily identical with Second-Language-Teaching research. However, when making decisions teachers can still benefit from the knowledge of SLA research findings, though these findings may not be directly applicable to the classroom. When conducting needs assessment, they may also benefit from the use of some of the concepts and instruments developed in the process of research of SLA.

A critical evaluation of an SLA research technique for measuring attitudes and the implications for language education is presented by Stephen Gaies and Jacqueline Beebe. The research technique is called the Matched-Guise Technique. According to Richards, Platt and Weber (1985:171) as quoted, the matched-guise technique is the use of recorded voices of people speaking first in one dialect or language and then in another; that is two "guises" ... The recordings are played to listeners who do not know that the two samples of speech are from the same person and who judge the two guises of the same speaker as though they were judging two separate speakers.

It is hoped that by using the technique, the researcher will be able to elicit reactions to particular codes by having subjects respond to taped samples of those codes rather than by having subjects express opinions about the codes themselves; in addition, s/he will be able to control all variables other than the codes themselves.

As a conclusion, the writers argue that some of the advantages of the technique as it was originally designed may be more apparent than real, but it can provide useful information which can and needs to be confirmed by other means.

Rod Ellis presents a critical evaluation of the Interaction Hypothesis. He argues that as teaching can be viewed as interaction that provides learners with opportunities for learning, the study of the relationship between interaction and SLA will contribute significantly to pedagogy. He further suggests a revision of the Interaction Hypothesis since a theoretical account of how input made comprehensible through interactional modification results in acquisition is missing from the original hypothesis. In his revision he suggests that the process of acquiring an L2 involves three basic procedures, that is (1) noticing, (2) comparison, and (3) integration. He concludes his paper by advising caution in applying the results of SLA studies to language teaching. SLA research should serve as one of the means of illuminating language pedagogy, and teachers themselves should make their own decision in the light of their classroom experience.

*****

Eugenius Sadtono

SECOND LANGUAGE ACQUISITION RESEARCH IN THE LANGUAGE CLASSROOM

David Nunan

Abstract

The purpose of this paper is to take a critical look at the current state of SLA research, focusing in particular on research which is intended to provide directions for classroom teaching. The study is based on an analysis of 50 empirical investigations which have been carried out and reported in the literature over the last 25 years.

Introduction

Second language acquisition (SLA) refers to the ways in which any learner, child or adult learns a second or foreign language. The learning may take place in tutored or untutored environments, and in second or foreign language settings. The field of second language acquisition has undergone tremendous growth over the last twenty years. There are now journals devoted exclusively to the subject, and it is even possible to undertake Ph.D. programs in SLA.

SLA research is coming under increasing scrutiny by those involved in language pedagogy. Curriculum developers and materials writers are interested in insights which the research can provide into optimal sequencing orders for acquisition, while teachers are concerned with those pedagogic tasks which can best facilitate acquisition in the classroom.

For many language teachers, second language acquisition is synonymous with Krashen (1981, 1982) whose work contains strong claims of relevance to pedagogy. In recent years, however, this work has come under attack from a variety of sources. Perhaps the most trenchant (and, as yet, unanswered) critic is Gregg (1986) who, in a celebrated review wrote:

> There is a Monty Python review in which a radio interviewer tries to get Miss Ann Elk, a dinosaur expert, to explain her new theory about the brontosaurus. After a great deal of hemming and hawing, false starts and

> general time wasting, we are finally told this: Brontosauruses were very thin in the front, much, much thicker in the middle and then very thin again at the end. Most of us would agree that, as a theory, this is rather unsatisfactory (indeed, the interviewer shoots Miss Elk before she can tell us her second theory). But then it was not meant to be taken seriously as a theory.
>
> Reading The Input Hypothesis, which evidently is meant to be taken seriously, brings Miss Elk to mind. The Input Hypothesis is the latest in a series of books and articles in which Krashen pretty much repeats what he has said in all his other books and articles; that is, he offers 'what I call, perhaps audaciously, a theory of a second language acquisition' (p.vii). (There are perhaps more fitting words than 'audaciously'; and in fact Krashen usually drops the article and talks simply of second language acquisition theory, a location that makes the complex error of suggesting that his theory is a theory, that a second language theory exists, and that his theory is it.) (Gregg 1986: 116-7).

While many SLA researchers are cautious about deriving pedagogical implications from their research, others are not so coy. Given the controversies and disagreements among the researchers themselves, to what extent should the practitioner take these claims at their value? This question was one which preoccupied me to the point that, when I was invited to carry out a review of research for a special issue of *Studies in Second Language Acquisition*, I decided to use the question as my point of departure. I shall summarise this research later in the paper.

Background

In an important review of SLA research, Lightbown suggests that SLA studies "are designed to investigate questions about learners' use of their second language and processes which underly second language acquisition and use" (Lightbown 1985: 173). Her statement reflects a distinction which is commonly drawn between the product-oriented tradition in SLA research and the process-oriented tradition. The aim of product-oriented research is to describe and predict the stages through which learners pass in acquiring a second language. Process-oriented research, on the other hand, is aimed at identifying those pedagogic variables which may facilitate or impede acquisition.

The variables may relate to the learner, the teacher, the instructional treatment/environment or some form of interaction between these.

Questions which classroom SLA research seeks to address include the following:
What types of classroom organisation and grouping patterns facilitate second language development?
What task and activity types facilitate acquisition?
What are the characteristics of teacher talk (including questions, amount, error feedback, instructions, directions) and what are the implications of this talk for acquisition?
Does formal instruction make a difference to the rate and/or route of acquisition?
What affective variables correlate with second language achievement?
What type of input facilitates comprehension and, by implication, acquisition?
What interactional modifications facilitate comprehension and, by implication, acquisition?

One major strand of SLA research is that which has focused on similarities and differences between input and interaction inside and outside of the classroom. It has been observed that there are clear differences in both classroom and naturalistic settings in terms of patterns of interaction, language functions, types of teacher questions and so on. (See, for example, Pica 1983; Larsen-Freeman and Long, 1991 for a summary of similarities and differences between the two settings and the possible consequences of these for acquisition.) The implications of these differences, and the extent to which classroom interaction should resemble real life interaction, are still being debated (van Lier 1988).

Long (1983; 1986) has argued that there is no evidence in the SLA research literature that classroom instruction can alter the order in which learners acquire particular morphosyntactic feature of the target language, and in fact, all the evidence is to the contrary. However, his review of the literature suggests that formal instruction does seem to be advantageous in three areas. In the first instance, formal instruction can facilitate processes of acquisition. Focusing learners on form can help make target features salient. If instructed learners are compared with those picking up the language naturalistically, it is found that in the early stages of acquisition the instructed learners are, paradoxically, making more errors. In particular, they make 'oversuppliance' errors, inserting morphosyntactic items where they are not required, saying things such as 'He loving a dog' or 'The girls loves her dog'.

> Initially, it looks like a disaster for teachers that one of the effects of their work is that learners make new kinds of errors that even naturalistic learners don't make, but in the long run they seem to slide back down to the appropriate suppliance of these things over time. Whereas if you look at naturalistic learners much later in their development, they are still doing an awful lot of omission and deletion. So I think the processes are affected in a healthy way by formal instruction. (Long 1987:295)

In addition to facilitating acquisition processes, Long argues that instructed learners develop their skills much more quickly than uninstructed learners. Finally, in terms of ultimate attainment, it seems that instructed learners end up being much more proficient than those who do not receive formal instruction.

The Study

As I indicated at the beginning of this paper, I was recently invited to survey the current state of classroom-oriented research in second language acquisition. This study was carried out to subject to critical scrutiny some of the claims made in the literature. Fifty studies which purport to be relevant to teachers, curriculum developers and teachers were analysed according to five analytical dimensions. There were:

1. The environment in which the data were collected (Were they collected in classrooms, in naturalistic environments outside the classroom, in simulated settings, or in laboratories?)

2. The rationale of the research (Were the studies carried out principally to inform those concern with pedagogy?)

3. The research design and method of collection (Were the data collected through experiments or not? What methods were used?)

4. The type of data collected (Did the study yield quantitative or qualitative data?)

5 The type of analysis conducted (What grammatical, statistical, and interpretive analyses were carried out?)

Results

(For a detailed description and analysis of the survey, see *Studies in Second Language Acquisition*, 1991, 13, 2.)

Environment

The most surprising outcome of the survey was the fact that, of the fifty studies, only fifteen were actually carried out in language classrooms which were constituted for the purposes of teaching and learning rather than research. Twenty eight of the studies were carried out in laboratory, simulated or naturalistic environments, and seven studies were carried out in mixed environments (that is, some of the data were collected in classroom, and some in non-classroom environments.)

Design

Design refers to whether or not the data were collected through some form of experimentation or not. Not surprisingly, only two of the fifteen classroom-based studies utilised some form of experiment. Roughly half (thirteen of the twenty-eight) non-classroom studies involved an experiment, as did the studies carried out in 'mixed' environments (three out of the seven).

Method

There were eight methods used in the studies. These included observation, transcript, diary, elicitation, interview, introspection, questionnaire, and case study.

Elicitation was the most frequently employed data collection method, with exactly half of the studies using some form of elicitation procedure to obtain their data. Classified under this heading are studies which obtain their data by means of a stimulus, such as a picture, diagram, standardized test, etc. The use of such devices has been a feature of SLA research since the original morpheme order studies obtained data through the use of the Bilingual Syntax Measure. When evaluating research utilizing such devices, it is important to consider the extent to which the results obtained are an artifact of the elicitation devices employed (see,

for example, Nunan 1987 for a discussion on the dangers of deriving implications for SLA from standardised test data). One needs to be particularly cautious in making claims about acquisition orders, based on elicited data, as Ellis (1985) has pointed out. (See also, Eisenstein, Bailey and Madden, 1982).

'Observation' can be either focused or unfocused. Focused observation refers to studies in which the investigator looks for specific aspects of language and behaviour, usually with the assistance of an observational instrument for classifying the behaviour being investigated. From the tables, it can be seen that while non-classroom investigations tended to utilise some form of elicitation, classroom studies were more likely to utilise observation or transcript analysis. ('Transcripts' refers to the analysis of interactions which are not subjected initially to some form of categorisation, but which undergo interpretative analysis later.)

'Questionnaires' are defined as instruments in which prespecified information is collected from informants through either written or oral responses. Questionnaires can be either closed or relatively open-ended. A closed questionnaire solicits data which can be readily quantified (e.g. those which require subjects to circle the appropriate response), while an open questionnaire enables subjects to provide a free-form response. Constructing questionnaires which unambiguously elicit accurate responses is difficult, and questionnaires designed to obtain information about language learning have the additional complication of sometimes being mediated through the learner's first language.

Interviews can also be relatively closed or open-ended. According to their purpose, they may be conducted either in the learner's first or target language.

The term 'diary' is used as a form of shorthand to refer to written, discursive accounts of teaching or learning, and which therefore contain free-form accounts of the learning/teaching process. They may be kept by learners, teachers or outside researcher/observer.

The use of introspective methods has a long history in cognitive psychology (see, for example, the use of verbal reports and protocol analysis in Ericsson and Simon 1980; 1984), but has only recently made its appearance in second language research. The emerging status of the method is reviewed in Faerch and Kasper (1987).

From the study it was found that questionnaires, interviews, diaries and introspection are infrequently used in classroom-oriented SLA research. This may reflect the suspicion with which SLA researchers view introspective and self-report data.

Type of Data

Initially, it had been intended to classify the studies according to the type of data collected. However, it was found that virtually all of the studies were based on qualitative data as defined by Grotjahn, being either discursive (e.g. transcripts) or nominal (e.g. functional categories, observation schemes etc.) This analysis therefore did not add anything particularly useful or insightful and was discarded. (In the event, it may have been better to use the common-sense notion of numericality to define 'quantitative' rather Grotjahn's more stringent notion of whether or not the data are measured on an interval scale.)

Type of Analysis

Three types of analysis, linguistic, statistical and interpretative were conducted, and the results of this analysis are set out in Tables 1, 2 and 3.

Table 1
Classroom-based Studies: Type of Analysis

Linguistic	Functions	5
	Complexity	4
	Morphosyntax	3
	Suprasegmentals	1
	Lexis	1
Statistical	Correlation	2
	Chi square	2
	t-test	1
	U test	1
Interpretive		9

Table 2
Laboratory, simulated and naturalistic studies: Type of Analysis

Linguistic	Morphosyntax	9
	Functions	8
	Complexity	2
	Quantity	2
Statistical	Correlation	7
	Chi square	5
	t-test	6
	Factor analysis	2
	F-ratio	2
	Cronbach's alpha	2
	ANOVA	2
	ANCOVA	1
Interpretive		6

Table 3
Mixed studies: Type of Analysis

Linguistic		6
	Morphosyntax	2
Statistical	Chi square	3
	t-test	2
	Correlation	1
Interpretive		3

Not surprisingly, classroom studies tend to be more interpretative and make a more limited use of statistics than non-classroom studies. This reinforces the picture which emerged in relation to the design and methods issues already discussed. It should be pointed out that the notion of 'interpretative analysis' is a relativistic one. All studies, even those employing a true experimental design and

utilising inferential statistics contain some form of interpretative analysis, even if this is little more than a footnote to the discussion of research outcomes.

In their analysis of statistical tests in applied linguistics, Teleni and Baldauf (1988) classify techniques as either basic, intermediate or advanced. Basic techniques include descriptive statistics, Pearson product-moment coefficient, Chi square, independent *t* - test, dependent *t* - test and one-way ANOVA. Applying their scheme to the studies analysed here, we see that the great majority of studies (29 out of 39) employ basic statistical tools. Many of the studies analysed in this review can be criticised on their research designs. There are also deficiencies in the manner in which they are reported. This is particularly true of experimental studies and those employing statistical analysis. Basic information, such as the number of subjects and whether or not they were randomised are frequently either not reported or buried away in the body of the report. There are also studies which violate assumptions underlying the statistical procedures employed. One particular problem is the analysis of group means through *t* - tests or ANOVA when the size is far too small for the analysis to be valid. (See also Chaudron 1988 for the critique of the use of statistics in classroom research.)

Discussion

What conclusions can we draw from this study? I believe that there are five points which need to be considered by SLA researchers who are motivated by a desire to provide directions for teachers, curriculum developers and materials writers. In the first instance, I believe there is a need for more contextualised research. Secondly, there needs to be an extension to the theoretical bases of the research agendas. Thirdly, it would be highly desirable to extend the range of research tools, techniques and methods, in particular, adopting methods which have proved to be valuable in content classroom research. While accepting the salience of the distinction between process and product-oriented research, I believe that this distinction needs to be re-evaluated. Finally, I would call for a more active role for the classroom practitioner in applied research, both through engagement with mainstream researchers in collaborative research projects, and also through action based investigations in their own professional contexts.

From the data, it can be argued that we need far more classroom-based, as opposed to classroom-oriented research. Further, we need research which investigates linguistic behaviour in context. Notwithstanding Labov's observer's

paradox, this means investigating real behaviour in real classrooms. van Lier puts the case for contextualised research in the following way:

> [Classroom] interaction consists of actions - verbal and otherwise - which are interdependent, i.e. they influence and are influenced by other actions. Pulling any one action, or a selection of them, out of this interdependence for the purposes of studying them, complicates rather than facilitates their description, just as a handshake cannot be adequately described, let alone adequately understood, by considering the actions of the two persons involved separately. ... [He goes on to say] The L2 classroom can be defined as the gathering, for a given period of time, of two or more persons (one of whom generally assumes the role of instructor) for the purposes of language learning. This is the setting of classroom research, the place where the data are found. I have argued before that, for CR to be possible, this setting must be intact, and not expressly set up for the purposes of research. [For van Lier] the central question that L2 classroom research addresses can be expressed as follows: How to identify, describe and relate, in intersubjective terms, actions and contributions of participants in the L2 classroom, in such a way what their significance for language learning can be understood.
>
> (van Lier 1988: 47)

Extending the theoretical bases of the research

I believe that there is some justification in extending the theoretical bases upon which much of the research rests. Many of the studies in this survey derive their theoretical rationale from transformational-generative grammar (although this is often more by implication than explicit acknowledgement.) The work of Krashen, particularly the 'comprehensible input hypothesis', has also been influential, although, once again, this is not always explicitly acknowledged.

In particular, it is worth looking to the development of a research agenda utilising alternative forms of analysis, such as that provided by systemic linguistics (Halliday 1985). Research based on systemic-functional grammar has provided rich insights into the development of oral and written language in first language classrooms. For example, it has demonstrated the value of explicit instruction in the generic structure of texts for the mastery of different types of written texts (see, for example, Christie 1985; Hammond 1990). Most of this work has been carried out

in first language classrooms, and it is worth extending this to second language contexts. (There is some evidence that this is in fact beginning to happen. See, for example, Mohan, forthcoming).

Extending the range of research tools, techniques and methods

There is also a need to extend the range of research tools, techniques and methods, adopting and adapting these where appropriate from content classroom research. (See Nunan 1989 for techniques, such as verbal protocols, stimulated recall and seating chart observation records, and for examples of their use in exploring language classrooms.) Allied to this is the desirability of using more than one instrument to obtain multiple perspectives on the phenomena under investigation. From the data presented in Tables 1, 2 and 3, it can be seen that only a handful of studies utilised more than two instruments.

One particularly underutilised method is the case study which, while it is associated in mainstream educational research with ethnographic research (see, for example, Barlett, Kemmis and Gillard, 1982), is a research 'hybrid' in that it can utilize data from a range of sources. In fact, a great deal of research in content classrooms is of this type. (The classic classroom-based case study is Smith and Geoffrey 1968, which drew data from a variety of sources, but principally from participant observation, non-participant observation, introspection and diaries.) The use of single case research of the type used extensively in speech pathology and human communication disorders is also worth looking at. In addition, as Larsen-Freeman and Long point out, there is no reason why SLA research might not utilize methods from either end of the methodological spectrum. They argue that longitudinal, naturalistic investigations could be supplemented by elicitation data.

A study utilizing such a hybrid approach is reported in Spada (1990). This investigation sought to determine (a) how different teachers interpreted theories of communicative language teaching in terms of their classroom practice, and (b) whether different classroom practices had any effect on learning outcomes. Three teachers and their intermediate "communicatively-based" ESL classes were used in the study. Each class was observed for five hours a day, once a week, over a six-week period. Students were given a battery of pre- and post-tests including the Comprehensive English Language Test and the Michigan Test of English Language Proficiency. The study utilized the COLT observation scheme as well as a qualitative analysis of classroom activity types. This indicated that one of the

classes, Class A, differed from the other two in a number of ways. The qualitative analysis confirmed the class differences, showing, for example, that class A spent twice as much time on form-based work than class C and triple the time spent by class B. To investigate whether these differences contributed differently to the learners L2 proficiency, pre- and post-treatment test scores were compared in an analysis of covariance. Among other things, results indicated that groups B and C improved their listening significantly more than group A, despite the fact that class A spent considerably more time in listening practice than the other classes. This research demonstrates the fact that qualitative observation and analysis were needed in order to interpret the quantitative results.

One of the persistent issues associated with research methods and paradigms concerns the status of knowledge. Despite observations on the complementarity of qualitative and quantitative research, there is a view that scientific 'truth' is arrived at through objective, value-free observation and inexorable logic. Medawar, in his provocative treatise on the scientific method, has this to say:

> ...we have been brought up to believe that scientific discovery turns upon the use of a method analogous to and of the same logical stature as deduction, namely the method of Induction - a logically mechanised process of thought which, starting from simple declarations of fact arising out of the evidence of the senses, can lead us with certainty to the truth of general laws. This would be an intellectually disabling belief if anyone actually believed it, and it is one for which John Stuart Mill's methodology of science must take most of the blame. The chief weakness of Milian induction was its failure to distinguish between the acts of mind involved in discovery and in proof. ... If we abandon the idea of induction and draw a clear distinction between having an idea and testing it or trying it out - it is as simple as that, though it can be put more grandly - then the antitheses I have been discussing fade away.
>
> (Medawar 1984: 31)

In the introduction to this paper, I suggested that research traditions were value neutral, that the issue or question should dictate the appropriate procedure. However, it may well be that a more basic, philosophical orientation will dictate which questions one considers worth asking in the first place.

Process-oriented versus product-oriented research

A commonly drawn distinction has been drawn between process- and product-oriented classroom research. Process-oriented studies focus on input and interaction, while product studies focus on the outcomes of instructional treatment. The great majority of studies in this survey were process-oriented, looking at such things as the negotiation of meaning prompted by different types of classroom task (see, for example, Doughty and Pica, 1986; Brock, 1986; Nunan, 1987a), and the comprehensibility of input as measured by standardised comprehension measures (see, for example, Chaudron and Richards 1986). Process-product studies which look at language gains which result from various forms of input were much less prominent in the data.

Most process-oriented studies are predicated on hypothesised relationships between various forms of input/interaction and acquisition and do not attempt to measure language gains. Doughty and Pica (1986), for example, established that two-way information gap tasks prompted significantly more modified interaction than one-way information gap tasks, and that small group tasks prompted significantly more modified interaction than teacher fronted tasks. From this they argue that group work:

> ...is eminently capable of providing students with opportunities to produce the target language and to modify interaction. In keeping with second language acquisition theory, such modified interaction is claimed to make input comprehensible to learners and to lead ultimately to successful classroom second language acquisition.
>
> (Doughty and Pica 1986: 322)

This type of research does not demonstrate (nor was it intended to demonstrate) a relationship between modified interaction and language gains. The hypothesised relationship is predicated on the assumption that the existence of interactional modifications ensures that the interaction is proceeding at a level which maximises the potential for comprehensible input.

Similarly, the studies by Brock (1986) and Nunan (1987a) indicate that the use of referential rather than display questions by teachers stimulate the production of longer and more complex responses by learners. However, they do not demonstrate that this actually fuels the acquisition process. (Long and Crookes 1986 did establish a link between the use of referential questions and experiential content

gains. However, the results did not reach statistical significance.) The study by Spada (1990), and other studies by some of Spada's colleagues in Canada are among the few to attempt to establish process-product links.

In addition, it can be argued that many of these so called process oriented studies are nothing of the sort, that in fact process is treated as product: instances of negotiation, wait time, foreigner talk etc., are bundled together and counted, the inference being the bigger the bundle the better! (van Lier personal communication).

Genuine process studies are difficult to find, although they are beginning to appear more frequently in the literature. One such study is that by Freeman (forthcoming). Freeman began with the question: How does the teacher define what can or cannot go on in her teaching - how are the boundaries of possibility constructed and negotiated through the talk and activity of the teacher's work? During the course of the investigation the focus shifted, and the question became "How are authority and control distributed, through pedagogy and interaction, to build a shared understanding of the subject in question (in this instance French as a foreign language)?

Freeman became a participant observer in a French as a foreign language classroom, and his data base included lesson transcripts, field notes and interviews with the teacher and students. The analysis consists of discursive and interpretive work on the data base. Freeman concludes from his investigation that:

> The process of evolving shared understanding of what to learn and how to learn it is at the heart of what makes [the teacher's] classes work. It takes place against the backdrop of constant social interaction ... and is intimately tied to sharing authority and control. [The teacher] has been able to make public the process of creating and internalizing the language precisely because she allows the talk and activity in her class to be largely self-regulated. Students come to control themselves in their interactions; that control goes hand-in-hand with authority over the language. Both involve the responsibility to an inner sense of rightness for appropriate behaviour and for accurate language use. This responsibility is individual and collective. [The teacher] is a resource for the language and a source for criteria and explanations of correctness. Likewise she is the source of activity in the classroom and a resource for successful accomplishment of that activity.

Freeman's study is a valuable addition to the literature on several counts. It highlights social and interpersonal aspects of language learning which are often randomised out of the language learning equation. It is also an example of a genuine process study. In addition, the shift in focus which occurred during the course of the investigation reflected an interaction between data and analysis, an interaction which is not untypical of qualitative research (Kirk and Miller 1986), but which would be considered 'unscientific' within a strict psychometric paradigm. Finally, the very questions it poses differ considerably from those generally posed by SLA research. .

A role for the teacher in classroom research

Finally, I should like to suggest that teachers themselves become more actively involved in the research process. The development of skills in observing and documenting classroom action and interaction, particularly if these foster the adoption of a research orientation by teachers to their classrooms, provides a powerful impetus to professional self-renewal. This is exemplified in the action research programs described in Nunan (1989). Such an orientation implies a particular role for the teacher. It is inconsistent with either the teacher as passive recipient of someone else's curriculum, or the notion of teaching as technology. the teacher researcher is one who is involved in the critical appraisal of ideas and the informed application of those ideas in the classroom. It is also at odds with the 'methods' approach to language teaching with its constant search for the one best way. The teacher researcher is less concerned with a search for the one best method than with the exploration of a number of variables in a range of classrooms with a diversity of learner types. Such exploration may, in fact, reveal that the complex mix of elements and processes results in variable outcomes and that what works in one classroom with a particular group of learners may not be as successful in a different classroom with different learner types.

While such exploration and analysis might add to our basic knowledge of language learning, such an ideal need not necessarily be the only or even the primary rationale for teacher research. It may be more realistic for teachers to recreate and test against the reality of their own classrooms claims for published research. The research literature which is surveyed in this study is a rich source of ideas on issues, methods and approaches, and many of these studies can stimulate the teacher to ask 'what might happen in my particular classrooms with my particular learners as a result of a particular intervention? While not wishing to

denigrate the value of the scientific method, nor to discount the care which many researchers take to guard against threats to internal and external validity, it is worth bearing in mind Cronbach's comment that, "When we give proper weight to local conditions, any generalization is a working hypothesis, not a conclusion". (1975: 125)

Conclusion

In this paper, I have taken a critical look at the current state of SLA research. While there have been considerable advances in the field, both methodological and substantive, there is room for improvement, and I have tried to indicate, in general terms, where these improvements might be made. In particular, I have suggested that future research would benefit from the informed incorporation into their design and execution of the following five key points:

1. The implementation of more contextualised research: that is, classroom-based, as opposed to classroom-oriented, research.

2. An extension of the theoretical bases of research agendas.

3. An extension of the range of research tool, techniques and methods, adopting and adapting these where appropriate from content classroom research.

4. A re-evaluation of the distinction between process-oriented and product-oriented research.

5. A more active role for classroom practitioners in applied research.

REFERENCES

ALLEN, J., M. Frohlich and N. Spada. 1984. The Communicative Orientation of Second Language Teaching. In J. Handscombe, R. Orem and B. Taylor (eds.) On TESOL '83. Washington DC: TESOL.

AVERY, P., S. Ehrlich and C. Yorio. 1985. Prosodic Domains in Foreigner Talk Discourse. In S. Gass and C. Madden (eds.) Input in Second Language Acquisition. Rowley Mass.: Newbury House.

BAILEY, K. 1983. Competitiveness and Anxiety in Adult Second Language Learning: Looking at and Through the Diary Studies. In H. Seliger and M.H. Long (eds.) Classroom Oriented Research in Second Language Acquisition. Rowley Mass.: Newbury House.

BAILEY, N., C. Madden and S. Krashen. 1974. Is there a 'Natural' Sequence in Adult Second Language Learning? Language Learning, 24, 2, 235-243.

BARTLETT, L.V., S. Kemmis and G. Gillard. (eds.) 1982. Perspectives on Case Study. Geelong Aust.: Deakin University Press.

BEEBE, L. (1983). Risk-taking and the Language Learner. In H. Seliger and M.H. Long (eds.) Classroom Oriented Research in Second Language Acquisition. Rowley Mass.: Newbury House.

BIALYSTOCK, E. 1983. Inferencing: Testing the "Hypothesis-Testing Hypothesis. In H. Seliger and M.H. Long (eds.) Classroom Oriented Research in Second Language Acquisition. Rowley Mass.: Newbury House.

BROCK, C. 1986. The Effect of Referential Questions on ESL Classroom Discourse. TESOL Quarterly, 20, 1, 47-59.

BROCK, C., G. Crookes, R. Day and M.H. Long. 1986. Differential Effects of Corrective Feedback in Native Speaker - Non-native Speaker Conversation. In R. Day (ed.) Talking to Learn: Conversation in Second Language Acquisition. Rowley Mass.: Newbury House.

CHAUDRON, C. 1983. Foreigner talk in the Classroom - an Aid to Learning? In H. Seliger and M.H. Long. (eds.) 1983. Classroom Oriented Research in Second Language Acquisition. Rowley Mass.: Newbury House.

*CHAUDRON, C. 1986a. **The Interaction of Quantitative and Qualitative Approaches to Research: A View of the Second Language Classroom.** TESOL Quarterly, 20, 4, 709-717.*

*CHAUDRON, C. 1986b. **Teachers' Priorities in Correcting Learners' Errors in French Immersion Classes.** In R. Day (ed.) Talking to Learn: Conversation in Second Language Acquisition. Rowley Mass.: Newbury House.*

*CHAUDRON, C. 1988. **Second Language Classrooms: Research on Teaching and Learning.** Cambridge: Cambridge University Press.*

*CHAUDRON, C. and J. Richards. 1986. **The Effect of Discourse Markers on the Comprehension of Lectures.** Applied Linguistics, 7, 2, 113-127.*

*CHRISTIE, F. 1985. **Language Education.** Geelong Australia: Deakin University Press.*

*DOUGHTY, C. and T. Pica. 1986. **"Information Gap" Task: Do They Facilitate Second Language Acquisition?** TESOL Quarterly, 20, 2, 305-325.*

*DUFF, P. 1986. **Another Look at Interlanguage Talk; Taking Task to Task.** In R. Day (ed.) Talking to Learn: Conversation in Second Language Acquisition. Rowley Mass.: Newbury House.*

*DULAY, H. and M. Burt. 1973. **Should We Teach Children Syntax?** Language Learning, 23, 245-258.*

*DULAY, H. and M. Burt. 1974. **Natural Sequences in Child Second Language Acquisition.** Language Learning, 24, 37-53.*

*ELLIS, R. 1982. **Discourse Processes in Classroom Second Language Development.** Unpublished Doctoral Dissertion, University of London. Cited in Ellis 1985.*

*ELLIS, R. 1985a. **Understanding Second Language Acquisition.** Oxford: Oxford University Press.*

*ELLIS, R. 1985b. **Teacher-pupil Interaction in Second Language Development.** In S. Gass and C. Madden (eds.) Input in Second Language Acquisition. Rowley Mass.: Newbury House.*

*ELLIS, R. (ed.) 1987. **Second Language Acquisition in Context.** London: Prentice Hall.*

*ERICSSON, K.A. and H.A. Simon. 1980. **Verbal Reports as Data.** Psychological Review, 87, 215-251.*

*ERICSSON, K.A. and H.A. Simon. 1984. **Protocol Analysis.** Cambridge Mass.: MIT Press/Bradford.*

FAERCH, C. and G. Kasper. (eds.) 1987. Introspection in Second Language Research. Clevedon Avon: Multilingual Matters.

FATHMAN, A. and L. Precup. 1983. Influences of Age and Setting on Second Language Oral Proficiency. In K. Bailey, M.H. Long and S. Peck (eds.) Second Language Acquisition Studies. Rowley Mass.: Newbury House.

FREEMAN, D. forthcoming. Collaboration and Co-constructing a Shared Understanding of a Second Language Classroom. In D. Nunan (ed.) Collaborative Language Learning and Teaching. Cambridge: Cambridge University Press.

GAIES, S. 1983. Learner Feedback: An Exploratory Study of its Role in the Second Language Classroom. In H. Selinger and M.H. Long (eds.) Classroom Oriented Research in Second Language Acquisition. Rowley Mass.: Newbury House.

GASS, S. and E. Varonis. 1985. Task Variation and Non-native/Non-native Negotiation of Meaning. In S. Gass and C. Madden (eds.) Input in Second Language Acquisition. Rowley Mass.: Newbury House.

GREGG, K. 1986. The Input Hypothesis: Issues and Implications. TESOL Quarterly, 20, 1.

GROTJAHN, R. 1987. On the Methodological Basis of Introspective Methods. In C. Faerch and G. Kasper (eds.) Introspection in Second Language Research. Clevedon Avon: Multilingual Matters.

HALLIDAY, M.A.K. 1985. An Introduction to Functional Grammar. London: Arnold.

HAMMOND, J. 1990. Grammar and Language Teaching. In D. Nunan (ed.) Working Papers in Applied Linguistics, Vol. 1. Sydney: Macquarie University.

HANSEN, L. 1983. The Acquisition and Forgetting of Hindi - Urdu Negation by English-speaking Children. In K. Bailey, M.H. Long and S. Peck (eds.) Second Language Acquisition Studies. Rowley Mass.: Newbury House.

HAWKINS, B. 1985. Is an "Appropriate Response" always so Appropriate? In S. Gass and C. Madden (eds.) Input in Second Language Acquisition. Rowley Mass.: Newbury House.

HEATH, S.B. 1983. Ways with Words. Cambridge: Cambridge University Press.

HENNING, G. 1986. Quantitative Methods in Language Acquisition Research. TESOL Quarterly, 20, 4, 701-708.

HOLMES, J. 1986. Sex Differences and Language Use in the ESL Classroom. In B. Das (ed.) Communication and Learning in the Classroom Community. Singapore: RELC.

HUEBNER, T. 1983. A Longitudinal Analysis of the Acquisition of English. Ann Arbor: Karoma.

IOUP, G. 1983. Acquiring Complex Sentences in ESL. In K. Bailey, M.H. Long and S. Peck (eds.) Second Language Acquisition Studies. Rowley Mass.: Newbury House.

JOHNSTON, M. 1985. Syntactic and Morphological Progressions in Learner English. Canberra Australia: Department of Immigration and Ethnic Affairs.

JOHNSTON, M. 1987. Understanding Learner Language. In D. Nunan (ed.) Applying Second Language Acquisition Research. Kelley, P. 1983. The Question of Uniformity in Interlanguage Devel. pment. In K. Bailey, M.H. Long and S. Peck (eds.) Second Language Acquisition Studies. Rowley Mass.: Newbury House.

KIRK, J. and M. Miller. 1986. Reliability and Validity in Qualitative Research. Newbury Park Ca.: Sage Publications.

KLEIFGEN, J.A. 1985. Skilled Variation in a Kindergarten Teacher's Use of Foreigner Talk. In S. Gass and C. Madden (eds.) Input in Second Language Acquisition. Rowley Mass.: Newbury House.

LARSEN-FREEMAN, D. 1975. The Acquisition of Grammatical Morphemes by Adult ESL Students. TESOL Quarterly, 9, 409-430.

LARSEN-FREEMAN, D. and M.H. Long. 1991. An Introduction to Second Language Acquisition Research. London: Longman.

LECOMPTE, M.D. and J.P. Goetz. 1982. Problems of Reliability and Validity in Ethnographic Research. Review of Educational Research, 52, 1, 31-60.

LIGHTBOWN, P. 1983. Exploring Relationships between Developmental and Instructional Sequences in L2 Acquisition. In H. Seliger and M.H. Long (eds.) Classroom Oriented Research in Second Language Acquisition. Rowley: Newbury House.

LIGHTBOWN, P. 1985. Great Expectations: Second Language Acquisition Research and Classroom Teaching, Applied Linguistics, 6, 2, 173-189.

LONG, M.H. 1980. Inside the "Black Box": Methodological Issues in Classroom Research on Language Learning. Language Learning, 30, 1, 1-42.

LONG, M. 1987. Second Language Acquisition and the Language Curriculum. Prospect, 2, 3, 293-300.

LONG, M.H. and G. Crookes. 1986. Intervention Points in Second Language Classroom Processes. Paper presented at the RELC Regional Seminar, Singapore, April, 1986.

MEDAWAR, P. 1984. Pluto's Republic. Oxford: Oxford University Press.

MOHAN, B. forthcoming. A Language Socialisation View of Cooperative Tasks. In D. Nunan (ed.) Collaborative Language Learning and Teaching. Cambridge: Cambridge University Press.

NUNAN, D. 1987a. Communicative Language Teaching: Making it work. ELT Journal, 41, 2.

NUNAN, D. (ed.) 1987b. Applying Second Language Acquisition Research. Adelaide: NCRC.

NUNAN, D. 1989. Understanding Language Classrooms. London: Prentice-Hall International.

NYSTROM, N.J. 1983. Teacher-student Interaction in Bilingual Classrooms: Four Approaches to Error Feedback. In H. Seliger and M.H. Long (eds.) Classroom Oriented Research in Second Language Acquisition. Rowley: Newbury House.

PARKER, K. and C. Chaudron. 1987. The Effects of Linguistics Simplification and Elaborative Modifications on L2 Comprehension. Paper presented at the 21st Annual TESOL Convention, Miami, April 1987.

PERKINS, K. and D. Larsen-Freeman. 1975. The Effect of Formal Language Instruction on the Order of Morpheme Acquisition. Language Learning, 25, 2, 237-243.

PIENEMANN, M. 1984. Psychological Constraints on the Teachability of Languages. SSLA, 6, 186-214.

PICA, T. 1983. The Role of Language Context in Second Language Acquisition. Interlanguage Studies Bulletin, 7, 101-123.

PICA, T. & M.H. Long. 1986. The Linguistic and Conversational Performance of Experiences and Inexperienced Teachers. In R. Day (ed.) Talking to Learn: Conversation in Second Language Acquisition. Rowley Mass.: Newbury House.

POLITZER, R. 1970. Some Reflections on 'Good' and 'Bad' Language Teaching Behaviours. Language Learning, 20, 1, 31-43.

PORTER, P. 1986. How Learners Talk to Each Other: Input and Interaction in Task-centred Discussions. In R. Day (ed.) Talking to Learn: Conversation in Second Language Acquisition. Rowley Mass.: Newbury House.

REICHARDT, C. and T. Cook. 1979. Beyond Qualitative and Quantitative Methods. In T. Cook and C. Reichardt (eds.) Qualitative and Quantitative Methods in Education Research. Beverley Hills: Sage Publications.

RULON, K. and J. McCreary. 1986. Negotiation of Content: Teacher-fronted and Small Group Interaction. In R. Day (ed.) Talking to Learn: Conversation in Second Language Acquisition. Rowley Mass.: Newbury House.

SATO, C. 1984. Phonological Processes in Second Language Acquisition: Another Look at Interlanguage Syllable Structure. Language Learning, 34, 4, 43-57.

SAVILLE-TROIKE, M. 1985. Cultural Input in Second Language Learning. In S. Gass and C. Madden (eds.) Input in Second Language Acquisition. Rowley Mass.: Newbury House.

SELIGER, H. 1983. Learner Interaction in the Classroom and its Effects on Language Acquisition. In H. Seliger and M.H. Long (eds.) Classroom Oriented Research in Second Language Acquisition. Rowley Mass.: Newbury House.

SELIGER, H. and M.H. Long. (eds.) 1983. Classroom Oriented Research in Second Language Acquisition. Rowley Mass.: Newbury House.

SCHINKE-LLANO, L. 1983. Foreigner Talk in Content Classrooms. In H. Seliger and M.H. Long. (eds.) 1983. Classroom Oriented Research in Second Language Acquisition. Rowley Mass.: Newbury House.

SCHINKE-LLANO, L. 1986. Foreigner Talk in Joint Cognitive Activities. In R. Day (ed.) Talking to Learn: Conversation in Second Language Acquisition. Rowley Mass.: Newbury House.

*SMITH, L.M. and W. Geoffrey. 1968. **The Complexities of an Urban Classroom.: An Analysis Toward a General Theory of Teaching.** New York: Holt, Rinehardt and Winston.*

*SNOW, M-A., J. Gavan, R.N. Campbell. 1983. **The Pilot Class of the Culver City Spanish Immersion Program: A follow-up report, or whatever happened to the immersion class of '78.** In K. Bailey, M.H. Long and S. Peck (eds.) Second Language Acquisition Studies. Rowley Mass.: Newbury House.*

*SPADA, N. 1990. **Observing Classroom Behaviours and Learning Outcomes in Different Second Language Programs.** In J. Richards and D. Nunan (eds.) Second Language Teacher Education. Cambridge: Cambridge University Press.*

*STRONG, M. 1986. **Teacher Language to Limited English Speakers in Bilingual and Submersion Classrooms.** In R. Day (ed.) Talking to Learn: Conversation in Second Language Acquisition. Rowley Mass.: Newbury House.*

*SWAIN, M. 1985. **Communicative Competence:** Some Roles of Comprehensible Input and Comprehensible Output in its Development. In S. Gass and C. Madden (eds.) Input in Second Language Acquisition. Rowley Mass.: Newbury House.*

*TELENI, V. and R.B. Baldauf. 1988. **Statistical Techniques Used in Three Applied Linguistics Journals: Language Learning, Applied Linguistics and TESOL Quarterly, 1980-1986: Implications for Readers and Researchers.** Unpublished Ms. Towell, R. 1987. Variability and Progress in the Language Development of Advanced Learners of a Foreign Language. In R. Ellis (ed.) Second Language Acquisition in Context. London: Prentice-Hall International.*

*van LIER, L. 1988. **The Classroom and the Language Learner.** London: Longman.*

*WALLACE, C. 1987. **Variability and the Reading Behaviour of L2 Readers.** In R. Ellis (ed.) Second Language Acquisition in Context. London: Prentice-Hall International.*

*WESCHE, M.B. and D. Ready. 1985. **Foreigner Talk in the University Classroom.** In S. Gass and C. Madden (eds.) Input in Second Language Acquisition. Rowley Mass.: Newbury House.*

*WEINERT, R. 1987. **Processes in Classroom Second Language Development: The Acquisition of Negation in German.** In R. Ellis (ed.) Second Language Acquisition in Context. London: Prentice-Hall.*

*WILLING, K. 1988. **Learning Styles in Adult Migrant Education.** Adelaide, Australia: NCRC.*

WOOD, H. 1978. ***Developmental Sequences in Naturalistic L2 Acquisition.*** *In E. Hatch (ed.) Second Language Acquisition. Rowley Mass.: Newbury House.*

WOLFSON, N. 1986. ***Research Methodology and the Question of Validity.*** *TESOL Quarterly, 20, 4, 689-699.*

WONG-FILLMORE, L. 1985. ***When Does Teacher Talk Work as Input?*** *In S. Gass and C. Madden (eds.) Input in Second Language Acquisition. Rowley Mass.: Newbury House.*

ZUENGLER, J. 1985. ***Phonological Aspects of Input in NS-NNS Interactions.*** *In S. Gass and C. Madden (eds.) Input in Second Language Acquisition. Rowley Mass.: Newbury House.*

A PLACE FOR SECOND LANGUAGE ACQUISITION IN TEACHER DEVELOPMENT AND IN TEACHER EDUCATION PROGRAMMES

Rod Bolitho

Any discussion of the relevance of Second Language Acquisition studies to teacher training programmes needs to take place within the broader context of the uneasy relationship between theory and practice in general, and between Applied Linguistics and English language teaching in particular. This relationship has long been a matter of concern to those involved in teacher education. Various attempts to define it have appeared in recent years Brumfit and Rossner (1982) discuss it in terms of decision-making, Widdowson's (1984) view takes the form of a homily, and a teacher training perspective is offered in Bolitho (1987). Ramani (1987) offers a route which leads teachers from their own practices towards a relevant 'theory'. Teachers often take up extreme positions, either deferring totally to theory or rejecting it out of hand as irrelevant to classroom issues. A position which many trainers seem to have arrived at and found useful, is that teachers need to understand why certain things work or don't work in classrooms and why materials writers and syllabus designers take certain decisions. 'Theory' may provide a part of the answer to some of those questions, though we cannot always be sure. What is certain however, is that one very worthwhile aim in teacher education is to turn out 'principled practitioners': teachers capable of asking the right questions and keeping the answers they may get from theory in a robust perspective of their own. Such a perspective is all the more necessary given the confusion and disagreement about research findings and language teaching theories in recent years. All too often, as Krashen (1989) points out, theorists have 'failed to deliver' and have, in the process, lost the respect of language teachers.

Yet there are dangers there for all to see. Applied Linguistics is, in many parts of the world, establishing itself as a kind of parent discipline to language teaching. Many applied linguists are involved in the training of teachers of English, and the pursuit of linguistic theory and research is all too often seen as a higher order activity than teaching. The more readily teachers acquiesce in this implied power structure, the more likely it is that theorists and researchers will attempt to set the agenda for classroom practice.

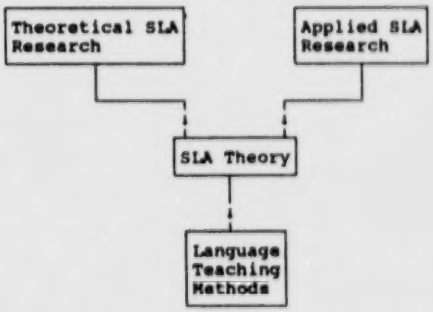
Conferences and seminars such as this one at RELC in 1991 can often contribute, unwittingly to the perpetuation of this kind of hierarchy. When researchers are given a platform to report their findings to an audience consisting largely of teachers, they carry a great burden of responsibility. It is all too easy to 'blind an audience with science' and to reveal research findings which are often derived from contexts quite alien to the majority of the listeners. The difficulties involved in carrying out large scale second language acquisition projects mean that many studies are limited to small groups of learners in well-defined and sometimes privileged learning situations. Teachers listen in awed silence to papers presented by researchers who are clearly expert in their own field, and the terms on which they can understand are defined by the speakers, not by the listeners. The 'code' which Second Language Acquisitionists have established to facilitate peer communication is not readily comprehensible to outsider groups, such as teachers. Very few teachers are versed in the methodology of research at any level. Yet we are all in the business of communication, and the onus is surely on researchers, as it is on specialists in any field, to find ways of describing their work to a lay audience with a legitimate interest in it. The reading of academic papers, for example, may be appropriate in a closed peer group, where everyone accepts it as a convention, and decoding presents no problems. In a conference where the audience is mixed, it simply results in miscommunication or even total alienation. When teachers and theorists or researchers meet, it should be an opportunity for genuine dialogue between professionals of equal standing. Teachers should not come away from such an encounter feeling guilty (about what they don't know), belittled, alienated or devalued. A decision to attend a conference or seminar is, after all, usually not taken lightly, and underlying it is the expectation that the event will contribute to one's professional development, though it may be wrong to expect solutions to problems or classroom recipes from such events. In this connection teachers and trainers will do well to remember that language teaching has a documented history stretching back, according to Kelly (1969) for 25 centuries whereas Second Language Acquisition has been recognised as a discipline for barely 25 years (though there have, of course, been many theories of language acquisition over the centuries). Surely such accumulated classroom experience is worth <u>something!</u>

Teacher educators, too, have a perspective on seminars of this sort, and on theories of Second Language Acquisition. This is a field which has had a considerable impact on language teaching in recent years, largely through the popular appeal of Krashen's ideas. His 'input hypothesis' and 'monitor model', allied to his views on learners' errors presented persuasively not only in print (Krashen 1982) but also on television, to mass audiences, demanded our attention

since they seemed to have obvious implications for classroom practice. Many teachers found his notions of 'comprehensible input' and errors as 'stepping stones on the way to learning' to be relevant and attractive. Indeed, these ideas, partly realised in Asher's 'Total Physical Response' Method, described in Asher (1969), formed the basis of a major national teacher training project in Indonesia (see Tomlinson 1990 for a full account). The prominence accorded to Krashen's ideas understandably led to criticism, too. Not everyone was so easily persuaded, and intuitive doubts were expressed at a very early stage (see , for example, Lowe 1983), to be followed later by more carefully elaborated positions (see, for example, Ellis 1985). Most of the objections are to the speculative nature of Krashen's ideas, which are not grounded in research. Those concerned with the education or training of language teachers need to decide how these arguments and counter-arguments can best be presented to teachers and trainee teachers who are primarily concerned with classroom-level decisions. In short, they have to decide how and in what measure to refer to Second Language Acquisition in teacher education programmes.

In a later book, Krashen (1989) refers to the relationship between research and practice with the help of a diagram which is reproduced here (Fig. 1)

Figure 1



from Krashen (1989:46)

He argues the case for the theorist (himself, for example!) as a mediator between research findings and classroom practice. (It is fair to state that he recognises there are many other areas of enquiry which inform language teaching apart from Second Language Acquisition, and that teachers' own insights and intuitions are of value.) But his diagram is based on a 'top-down' view (look at the arrows!) and takes no obvious account of the ways in which teacher educators also have to mediate between research findings and theory on the one hand, and classroom practice on the other. Those other areas of enquiry (linguistics, lexicography, humanistic psychology etc) <u>are</u> important to language teacher educators, and Second Language Acquisition has to take its place among them on a crowded teacher education syllabus. In the light of this, a legitimate question is: which areas of Second Language Acquisition should we focus on in our programmes, where time is so limited? For the time being the following areas seem relevant to both initial and in-service programmes, partly because of the attention they have attracted, and partly because they have clearly identifiable practical implications:

- interlanguage and errors
- learning vs acquisition
- learning styles and learning strategies (though Second Language Acquisition research is only one piece in the jigsaw here)

In addition, on some in-service programmes, where the focus is on continuing professional development, it seems appropriate to broach the issue of classroom research, though here again, Second Language Acquisition is only one possible area on which to focus attention.

Having decided which aspects of Second Language Acquisition to include in the course, the teacher educator needs to decide <u>how</u> to present them and (in many cases) <u>who</u> should present them. Taking the latter question first, and given the problems outlined earlier in this paper over teachers' relationship with research, I would like to advance the view that those engaged in research are not always best equipped to teach in their own discipline area on training courses. It may seem exciting to study in a research-oriented unit where frontiers are being pushed back, but it is not necessarily healthy for teachers and trainees. Indeed, there are many good reasons why the language teaching profession should not react too quickly to research findings. Changes in syllabus, materials and methods, if perceived as being top-down and too frequent, cause difficulty and even distress for learners and teacher alike. The natural 'home' of Second Language Acquisition research work is

in Departments of Linguistics or Psychology, whereas most teacher training is rightly carried out in Faculties or Colleges of Education.

It is, however, all too easy, in many institutions, for those running training courses to assume that the only way to deal with Second Language Acquisition is to 'buy in' the services of a specialist who may then find it difficult to present the subject in an accessible way. The alternative is for the trainer (whose main Language Acquisition studies into the course. This places the onus on trainers to keep reasonably up to date with research findings through the literature, and to interpret them for the purposes of their trainees. Trainers may also wish to provide their trainees with a basic grounding in classroom research, in order to empower them to conduct their own investigations when the need arises. I believe this mediating role to be vitally important if practitioners' 'blocks' about research are to be overcome, and if practical concerns are to be successfully articulated to researchers.

The other question ('how to deal with a Second Language Acquisition component on a training course') is then rather easier to answer. Many trainers are aware of the value of building on the existing experience of their trainees as a starting point from which theoretical issues can be approached. Indeed, it has often been stated that most teachers and learners have their own 'theory' of language learning, usually more implicit than explicit. Part of a trainer's responsibility is to encourage trainees to articulate this 'theory'.

Using this as a basis, trainers can consider an approach which integrates SLA work into an experiential learning cycle such as the one illustrated in Fig. 2., derived from models proposed by Gibbs and Habeshaw (1989) and Dennison and Kirk (1990) (though there are other similar models elsewhere in the literature of pedagogy).

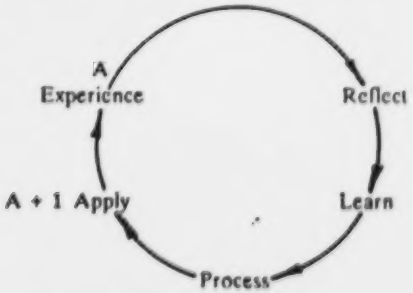


Fig. 2

In an INSET session on Errors and Interlanguage, for example, the trainer might start by asking teachers to describe their attitudes to learners' errors and their current classroom practices. Teachers could be encouraged to compare their practices with those of colleagues (i.e. reflecting on their own experience). The trainer would then summarise key issues and unresolved questions before asking trainees to read a couple of accessible extracts from teachers' handbooks, e.g. Edge (1989) or Corder (1981) or from background literature, e.g. Krashen (1982) (the learning stage). Teachers could then be asked to reassess their own attitudes and practices in the light of this new input (the 'processing' stage, much of which may happen after the course), thereby arriving at a new position which they will go on to apply in their own classrooms. So they will have absorbed useful insights from SLA which will have played a part in moving them from their original position (A in the diagram) to a revised position (A + 1 in the diagram). This may entail a revised attitude to the treatment of learners' errors based on an enhanced understanding of their status and significance. In arriving at this new position, they have had their own views heard and respected, and have been encouraged not to abandon them completely but to modify them where appropriate in the light of the interpreted research findings. Appendix One consists of a training sequence based on this approach. On a pre-service course, a similar approach could be taken, using the trainees' views of error as language learners as the experiential starting point. In either case, the model allows theory and research findings to be assimilated digestibly into the overall methodology of training courses and to be kept in sensible perspective.

If trainers are to play this kind of mediating role successfully, and if teachers and researchers are to develop a healthy and sensible working relationship, certain conditions will have to be met.

1. To start with, Krashen's diagram could usefully be modified (as in fig. 3) to imply two-way dialogue rather than one-way transmission, and to take account of insights from other fields. Researchers often call for co-operation from teachers, whose learners are needed as subjects of research, and they frequently urge teachers to give attention to their findings. They have no particular right to expect this of teachers unless they are prepared to listen as well. It is as legitimate for teachers to make demands on researchers (e.g. by helping to establish a research agenda) as it is for researchers to influence what goes on in classrooms. Teacher trainers, many of whom spend a fair amount of time observing in classrooms, may have a useful perspective to offer here, too. An example might be useful.

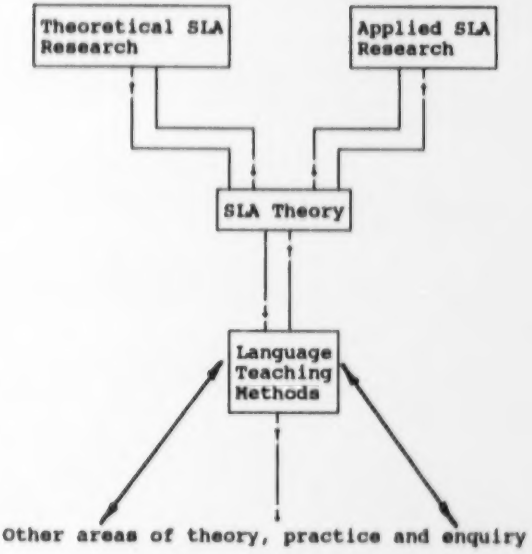
Teachers, particularly in a region like S.E. Asia, understandably get weary of hearing the results of small-scale SLA studies carried out in classroom contexts which are almost totally unrecognisable to them. Is it really unreasonable to ask researchers to turn their attention to larger scale studies in underprivileged classrooms?

2. Successful two-way communication depends on mutual comprehensibility. Most teachers talk in terms that researchers can readily understand. Many researchers have become used to talking in terms which are only comprehensible to other researchers. The register of research is remote and inaccessible to many teachers. The onus here is on researchers. When invited to address conference and seminar audiences consisting largely of teachers, they need to present their ideas in an accessible way, in terms which will make sense to teachers. It simply will not do to give the same paper as they gave at the last specialist SLA conference in the same way.

3. Following on from this, there is a point to be made about literature. Comparatively few books on SLA are written in terms that teachers can readily understand. Krashen (1989), Ellis (1985) and, with some reservations, Littlewood (1984) are notable exceptions. In such a fast-developing area (in which we hear that Krashen's theories, for example, are already 'old hat'), there is a need for regular 'state-of-the-art' publication in non-intimidating language, to allow all those with a legitimate interest in ideas from research to assess them on their own terms. If researchers wish to be taken seriously outside their limited circuit, they will have to take responsibility for producing this kind of intermediate literature, which teacher trainers need if they are to deal with SLA successfully on their courses.

4. There is a crying need for more teacher training material in the field of SLA. Selinker and Gass (1984), now sadly out of print, is an example of immediately usable training material on the form of awareness-raising tasks based on samples of learner language. Trainers need banks of this type of task-based material, both for class use and for self-access purposes. A fruitful joint project for a trainer and a researcher, perhaps?

Figure 3



If our aim in teacher education is to train our trainees, pre-service and in-service, to be 'principled practitioners', we need to help them to ask the right questions, to arrive at a better understanding of the whys and wherefores of successful language learning, to resist the attractions of panaceas and recipe-type solutions, and to lay the foundations for continuing professional development. If SLA researchers are to claim a role on this valuable process, they must understand how best to play it. Trainers need principled support, not confusing messages from the world of research.

It has been the purpose of this paper to identify and discuss some of the causes of such confusion and to attempt to describe the kind of support which might be most useful.

REFERENCES

*ASHER, J. 1969. **The Total Physical Response Approach to Second Language Learning.** Modern Language Learning Journal 53.*

*BOLITHO, R. 1987. **Teaching, Teacher Training and Applied Linguistics.** In Bickley V. (ed.), Re-Exploring CELT (ILE, Hong Kong).*

*BRUMFIT, C., and R. Rossner. 1982. **The Decision Pyramid and Teacher Training for ELT.** ELT Journal 36/4.*

*CORDER, S.P. 1981. **Error Analysis and Interlanguage.** OUP.*

*DENNISON, B. and R. Kirk. 1990. **Do, Review, Learn, Apply.** Basil Blackwell.*

*EDGE, J. 1989. **Mistakes and Correction.** Longman.*

*ELLIS, R. 1985. **Understanding Second Language Acquisition** OUP.*

*GIBBS, G. and T. Habeshaw. 1989. **Preparing to Teach.** Technical and Educational Services Ltd.*

*JOHNSON, K. 1988. **Mistake Correction.** ELT Journal 42/2.*

*KELLY, L.G. 1969. **25 Centuries of Language Teaching.** Newbury House.*

*KRASHEN, S.D. 1982. **Second Language Acquisition and Second Language Learning.** Pergamon.*

*KRASHEN, S.D. 1989. **Language Acquisition and Language Education.** Prentice Hall.*

*LOWE, M. 1983. **Language Acquisition and Language Learning: A Reappraisal.** In Holden S. (ed) Focus on the Learner. Modern English Publications/British Council.*

*MURPHY, D. 1986. **Communication and Correction in the Classroom.** ELT Journal 40/2.*

*RAMANI, E. 1987. **Theorizing from the Classroom.** ELT Journal 41/1.*

*SELINKER, L. and S. Gass. 1984. **A Workbook in Second Language Acquisition.** Newbury House.*

*TOMLINSON, B.J. 1990. **Managing Change in Indonesian High Schools.** ELT Journal 44/1.*

*WIDDOWSON, H.G. 1984. **The Incentive Value of Theory in Teacher Education.** ELT Journal 38/2.*

Appendix

Sample In-Service Training Material

LEARNERS' ERRORS

1. What do you do about errors in your learners' spoken English? Tick any box(es) that apply to you.

I correct them all	☐
I correct some of them	☐
I ignore them	☐
I just point them out	☐
I get my learners to correct each other	☐
I get them to self-correct whenever possible	☐
I never let them make any errors in the first place	☐
Other strategies (please specify)	☐

2. How do you feel about errors in your learners' spoken English? Tick any box(es) that apply to you.

I get irritated by my learners' errors	☐
I feel despondent about all the errors they make	☐
I am sometimes confused by their errors: there seems to be no pattern	☐
I feel responsible for their errors	☐
It's my duty as a teacher to correct errors	☐
I like correcting my learner's errors	☐
I dislike correcting my learner's errors	☐
I'm afraid that errors which are left uncorrected will be passed on to other learners	☐
Other feelings (please specify)	☐

3. Look at this dialogue written as a composition exercise by a learner. Correct it, ready to return, as if she were one of your own students.

A DIALOGUE BETWEEN JUMA AND JOHN

IN A TELEPHONE

Juma:	Hallo!
Telephone operator:	Hallo! What do you want me to do for you?
Juma:	I want to talk to John please.
Telephone operator:	What is your name?
Juma:	I am Juma Ali.
Telephone operator:	Ah! Juma! Okay wait for not more than thirty minutes.
John:	Hallo! dear friend.
Juma:	Hallo John. How are you?
John:	I am quite fine.
Juma:	I hope you are not surprised for a call at this time. I have good news for you but first will you be free this evening.
John:	I don't think.
Juma:	I have a very urgent message for you so because you will not be free don't dream of getting that chance again.
John:	Oh! no Juma. Yes I have gotten an idea I will ask my friend to replace me for sometime.
Juma:	Good. That is a good and a reasonable idea. I love that. That is why I love you John.
Both:	Ha! ha! ha! (Both laughed simultaneously).

Now work in groups of 4 and discuss your various approaches to correcting this piece of work. Note down any interesting differences of opinion to report back to the whole group. Stay in these groups of 4 for steps 4-6 which now follow.

4. How do you think learners feel about their errors:

   a) in spoken English?
   b) in written English?

5. What do your learners do about their errors

   a) in spoken English?
   b) in written English?

6. What significance do errors have in the educational system of your country

   a) for teachers?
   b) for learners?
   c) for learners' parents?
   d) in examinations?

7. Follow-up

   The Course members now divide into two groups (A +B)

   Everyone in Group A receives a copy of 2 extracts from Krashen (1982) pp 74-76 and pp 116-119.

   Everyone in Group B receives an extract from Edge (1989) pp 1-17.

   Working in pairs within each group, prepare 5 key reading comprehension questions on each extract. When you've done this, get together with a pair from the other group and exchange your texts and questions. Try out and comment on the questions you have been set.

8. Further Reading

   Johnson's 1988 article
   Murphy's 1986 article
   The rest of Edge's book
   The rest of Krashen's book
   Corder (1981) if you'd really like to know more about Error Analysis

DIMENSIONS IN THE ACQUISITION OF ORAL LANGUAGE

Martin Bygate and Don Porter

Introduction

This paper is a preliminary exploration of two questions: what is the effect of task familiarity on language production; and, more generally, in what ways does oral production on a given task improve? It reports a small pilot study for a larger investigation as part of the Oral Language Proficiency Project under way at Reading.

Curiously, although there has been enormous expansion in second language acquisition studies over the last twenty years, little attention has been paid to task-based development. Research in the seventies concentrated largely either on ethnographic studies aimed at understanding communicative competence, or on understanding child language acquisition. Since the early eighties, a major development in the context of language teaching and applied language studies has been to explore the nature of language use on specific tasks (cf. Brown & Yule, 1983; Bygate, 1987, 1988; Faerch & Kasper, 1983; Long & Porter, 1985). While the focus of these studies (and others) has quite rightly emphasised the patterns of language use, most writers have also recognised the need to widen the focus to study the nature of task-based development (eg. Long, 1989), since what matters most to us as teachers is the relationship between task and language development. This relationship is the focus of the present paper.

1.1 Previous Studies

Child language acquisition studies have already to some extent pursued this line. Snow (1987) studied the use of language in typical contexts (of explanation and definition) while Bruner (1983) reported a major project on the role of recurrent context in early child language acquisition. This field constitutes a rich variety of directions to explore in second language acquisition. The present paper limits itself to a consideration of the nature of improved performance on a given task, and the

extent to which this can be attributed to the familiar "task-practice" effect. Improvement here will be investigated on two dimensions - quantitative formal measures; and a small number of measures of fluency.

Various writers have explored characteristics of oral production on given tasks, although not from a developmental point of view. Long & Porter (1985) reviewed studies exploring the relationship between the incidence of negotiation of meaning and task type. Anderson (1985) reported a study in which task type influenced both listener involvement and speakers' performance. Faerch & Kasper (1983) included a range of studies focusing on L2 performance in which tasks are the independent variable. Most of those studies however were attempting to develop an inventory of strategies used in an arbitrary range of tasks. In a more controlled study, Poulisse & Schils (1989) reported a task effect in the use of communication strategies.

The negotiation of meaning and communication strategies are not however the only areas of interest for task-based research. In L1 studies, Tannen (1980) reported differences between the way Greek and American speakers carry out a narrative task. Using the same task as Tannen, Chafe (1980) analysed linguistic features of fluent production of L1 spoken discourse.

Such work has been less frequent in L2 studies. However, Dechert (1983) reported the performance of a non-native speaker describing a picture story from the point of view of the speaker's fluency. Lennon (1990) provides a rare developmental study, specifically of non-native speaker fluency, in which however the task type is uncontrolled conversational interaction. It could be argued that such studies are of potential use for an understanding of the dimensions of the development of fluency, but they are nevertheless some distance from applicability in teacher intervention. It is because tasks are a key element of teacher intervention (see Long, 1989; Nunan, 1989) that it seems of particular interest to observe how language performance develops, with the task being the controlled variable.

1.2 Effects of Task Familiarity

A well-known paper by Goldmann-Eisler (1961) reported the effect on pausing of the preplanning of a story-telling task. Goldmann-Eisler was able to show reduction in the amount of pausing as speakers were allowed more planning time. A more general argument concerning the nature of what might be called the practice effect has been proposed by Elinor Ochs (1979). The distinction between spoken

and written language, she argues, is less illuminating than that between planned and unplanned speech. the characteristics of planned speech are closer to those of written text, with fewer editing features and, as Goldmann-Eisler showed, probably with less pausing. It could therefore be argued that one of the factors promoting proficient performance (which would include fluency) would be familiarity with the task.

The argument would illuminate the foreign language learner's experience in various ways. For one thing, familiarity with a task would mean that a learner would have less work to do in planning the message. Familiarity with the information to be communicated would reduce the work involved in planning individual messages. And familiarity with its organisation would in principle improve overall planning and execution. The effect would be to improve performance.

There may be two major effects of task familiarity. One effect might be greater fluency. Alternatively, if fluency were not gained, there might be a gain in accuracy or in lexico-grammatical range. The question then is, how does performance improve, and does performance on a given task improve in the same way as performance on other tasks at the same point in time? Answers to questions of this sort could be expected to be of interest to teachers and testers, since such information would help to predict the kind of effect on performance of carrying out familiar rather than unfamiliar tasks. Teachers as well as testers could use this information in selecting tasks. This approach could also contribute to our understanding of language use. These then are the purposes of the study reported in this paper.

1.3 Focus on Analysis

Two important areas of interest in oral language performance are the language used, and the fluency of production. A number of measures of performance have been used in the past.

With respect to the language produced, development could be looked for in the overall organisation of the discourse (eg. narrative or descriptive structure, cf. Linde & Labov, 1975); in the syntactic units employed by speakers (eg. T-units or c-units, cf. Crookes, 1990; sentence complexity cf. Crystal, Fletcher & Garman, 1976); in the range of expressions and lexical items (range of vocabulary); in the incidence of

grammatical error; and in the ability to handle on-line problems of expression (communication strategies). A thorough attempt to understand language development would need to take account of all these dimensions. This report will limit its observations to a discussion of complexity in vocabulary range and complexity in clause-structure. The assumption would be that a person repeating a familiar task would be more likely to demonstrate more complex language: a wider range of vocabulary, more subordinate clauses, and greater clausal quantity than when doing an unfamiliar task.

Fluency can be identified through analysis of the incidence of pausing, and through the occurrence of repair (see Lennon, 1990, for a similar approach). It would be worth considering for a moment how fluency might be seen to develop with task familiarity.

Dechert (1983) suggests that speakers develop particular sequences of language which are uttered with greater fluency and less pausing, which he terms "islands of reliability". In his view, the speaker may have difficulty marshalling language to move from one area to another, and the transitions are marked by greater pausing and hesitation. Dechert's task is the one used in the present study.

It should be stressed here that it is in no sense our view that pausing is a mark of non-native speaker dysfluency and therefore something to be eradicated from learner language. Beattie (1980), examining native speaker recordings, suggested that in certain kinds of oral performance (tutorials), speakers will alternate between fluent and dysfluent passages of speech. However, the explanation provided by Beattie for the dysfluent sections is not that these indicate inadequate proficiency but rather the fact that at certain points in the discourse, speakers need to undertake general long term ("distal") message planning. Once the speaker has sorted out the direction of the message, fluent speech can be resumed. Thus, as Fulcher (1987) points out, hesitation and pausing is characteristic of native speaker talk. However, since pausing can be seen as a function of the planning load, the effect of task-familiarity on pausing in non-native speech is worth some study.

Indices of increased fluency may take different forms. For our present purposes we will limit our observations to the amount of pausing, and the amount of self-repair, the prediction being that both pausing and self-repair would decrease, at least up to a point, with increasing task familiarity (see Lennon 1990 for a contrasting finding).

2. The Study

A pilot study was set up at the University of Reading to examine the effects on performance of repeating a task. Non-native students were interviewed in Autumn 1990 (Time 1) and then again three months later (Time 2). Subjects were students of different language backgrounds who were studying on preparatory EAP courses at Reading. They were interviewed by experienced native-speaker interviewers. On both occasions the students were asked a range of general questions about their studies and reasons for choosing Reading as a suitable place for higher academic work, and were then asked by their interviewer to describe a short picture story used in Dechert 1983. On the second occasion, when the first story was familiar to them, the students were also asked to recount a new picture story. Interview sessions were audio- and video-recorded, and the recordings transcribed and analysed.

Comparisons were made between students' performance on story 1 at times 1 and 2, and between their performances on stories 1 and 2 at time 2. Four subjects were selected from a larger sample for this initial report. The purpose of the study then is to see whether any changes in the telling of the story 1 by time 2 are attributable to familiarity with the story, or whether those changes can also be found in the telling of story 2.

2.1 Units of Analysis

The transcripts were analysed using the following measures.

1. Individual pauses: filled, unfilled:
   instances of consecutive filled and unfilled pauses count as a single unit of dysfluency - the argument being that filling is simply a way of extending a pause; this will contribute to the total amount of pause time but will not reflect the number of decision points; pausing is therefore taken as an indication of the number of selection or access problems encountered by the speaker.

2. Repairs including:
   false starts, repetitions of words or utterances, incomplete fragments,redundant repeated words are counted individually - this reflects the speaker's uncertainty about the lexical decision; a repetition of a word may occur in order to maintain discourse coherence both for speaker and listener where the speaker has paused

and not found a more suitable lexical item; such a repetition may not have the same function as repetition without a preceding pause which may simply be providing thinking time before producing the following string; nonetheless pause+repetition can reasonably be interpreted as a mark of dysfluency since it signals the need to recover coherence in the interlocutor's short-term memory in the case of a word change, the dysfluency is in the speaker's slow and inaccurate lexical accessing. Whole utterances are counted as single units of dysfluency where they repeat or rephrase a message which was already expressed in an immediately preceding sequence of discourse (individual words are not counted, even though it could be that the longer the repetition the more covert planning time may be being created); repetitions are not taken into account where these occur at other points in the discourse with a different discourse function.

3. Vocabulary complexity:
   type: token ratio was calculated in this study using number of fluent words as the total number of tokens.

4. Syntactic complexity:
   syntactic complexity was gauged by calculating the ratios of:
   a) total number of words to the number of finite clauses;
   b) number of subordinate clauses to the number of main clauses;
   c) the total number of clauses.

2.2 Hypotheses

The Hypotheses were as follows:

H1 There would be fewer indices of dysfluency as identified under section 2.1.2 above in the retelling of story 1 than on the two tellings in unfamiliar mode (story 1 at time 1, and story 2)

H2 There would be a wider range of vocabulary items, and greater syntactic complexity as measured by incidence of main and subordinate finite clauses, in familiar mode than in unfamiliar mode.

3 Results

Results on the two sets of measures are presented in Tables 1 and 2.

Table 1: Comparison of fluency measures on the three tasks

		TASK 1(1)	TASK 1(2)	TASK 2
Ratio pauses:	S1	.40	.24	.22
no. of words	S2	.27	.11	.23
	S3	.18	.19	.10
Ratio repairs:	S1	.42	.17	.38
no. of words	S2	.13	.07	.12
	S3	.09	.09	.08
Total	S1	.82	.41	.60
	S2	.40	.18	.35
	S3	.27	.28	.18

Table 2: Comparison of complexity measures on the three tasks

		TASK 1(1)	TASK 1(2)	TASK 2
Vocab ratio	S1	.63	.65	.63
type:token	S2	.45	.34	.44
	S3	.36	.32	.31
Clause ratio	S1	6.0	6.00	6.55
no. words:	S2	.07	7.28	6.88
no. clauses	S3	7.60	6.73	6.00
Clause ratio	S1	.08	.00	.00
subord: main	S2	.08	.29	.14
clauses	S3	.07	.50	.12
No. of	S1	14	12	11
clauses	S2	15	22	16
	S3	15	20	21

3.1 Interpretation and Discussion

3.1.1 Hypothesis 1

Hypothesis 1 would predict lower ratios on Task 1(2) than in the other two columns in Table 1. The total figures do indeed tend to support this hypothesis for Students 1 and 2. Student 3 however goes against the trend. Her ratios are virtually the same for task 1(1) and 1(2), and both are higher than for task 2. That is to say, she is more fluent on task 2 with which she is unfamiliar. Student 3 may differ from the other two students in terms of general proficiency, and we might note that she is from Sri Lanka where English is widely used as a second language. The other two students are from EFL backgrounds. In their case, fluency is highest in the familiar mode (ie. task 1(2)).

Students 1 and 2 perform consistently more fluently on tasks 1(2) and 2 than on 1(1), both with respect to a decrease in repair and in pauses. While we can of course make only the most tentative of remarks about these results, they first suggest an overall improvement over the 3 month period. It is worth noting that in the case of Student 1, the improvement with respect to pausing appears to be generalised in the same strength from familiar to unfamiliar tasks. However, for Student 1 on pauses, and both Students 1 and 2 on repairs, the increase in fluency is very limited on the unfamiliar task. In other words, there is some evidence not only to support the hypothesis that fluency will be positively influenced by task familiarity, but also to suggest that learning will be potentiated by familiarity of task.

3.1.2 Hypothesis 2

A second way in which task familiarity could be expected to influence performance is in the increase of linguistic complexity. Hypothesis 2 allows for the fact that the degree of increase in complexity may be enhanced by task familiarity. It remains an open question whether fluency and complexity improve simultaneously. It could be that one would improve at the expense of the other.

Taking the results student by student, Student 2 improved in complexity on all measures on both familiar and unfamiliar tasks at time 2, with the exception of vocabulary type-token ratio. (Indeed one might note at this point that the type-

token measure of vocabulary range, does not reveal any notable trend in improvement for any of the students. This could be because no improvement took place in vocabulary range or because this is not a sufficiently sensitive measure for this purpose.) In addition, Student 2 showed a notable familiarity effect on task 1(2) which is consistent with hypothesis 2. This student appears to have improved therefore both in terms of fluency and linguistic complexity.

It is striking that Student 3 becomes more economical in terms of words over the three tasks, while at the same time improving in the number and complexity of clauses. A familiarity effect may have been responsible for the high incidence of subordinate clauses in task 1(2). The suggestion might be that this learner has improved principally in terms of syntactic performance.

Finally, with regard to Student 1, we note a lack of improvement in terms of linguistic complexity on all of the measures in table 2. This suggests that this student's improvement can be largely located in her increased fluency.

Conclusion

Two main suggestions could be made on the basis of this preliminary study. All relate to possible differences between learners. First of all, Student 3 seems largely immune to the effect of familiarity in all except clause subordination. The other two students on the other hand do seem to respond to the familiarity of the task.

Secondly, we would note that while Student 3 seems to improve mainly in terms of linguistic complexity, Student 1 seems to improve only in terms of fluency, and Student 2 seems to improve on both sets of measures. This suggests that there could be at least three different patterns of oral language development; in fluency, in linguistic complexity, or in both.

Various other variables will need consideration. For one thing, is there a ceiling effect on a given task; at what point does maximum fluency set in? A further point concerns the possible effects of more concentrated practice resembling the kind provided by teachers. And finally it remains to be seen what results will emerge from the study of a larger sample.

REFERENCES

ANDERSON, A. 1985. ***"What Can We Do to Promote Good Listening? An Experimental Search for One Possible Answer"*** *Paper given at BAAL Annual Meeting, Edinburg.*

BEATTIE, G.W. 1980. ***'Face to Face Interaction'.*** *In B. Butterworth (ed.): Language Production. Vol. 1: Speech and talk. London: Academic Press.*

BROWN, G & G. Yule. ***Teaching the Spoken Language.*** *Cambridge: CUP.*

BRUNER, J.S. 1983. ***Child's Talk.*** *Oxford: OUP.*

BYGATE, M. 1987. ***Speaking.*** *Oxford: OUP.*

BYGATE, M. 1988. ***"Units of Expression and Acquisition in Small Group Interaction"*** *Applied Linguistics. 9/1.*

CHAFE, W.L. 1980. ***The Pear Stories:*** *Cognitive, Cultural and Linguistic Aspects of Narrative Production. Norwood, NJ: Ablex.*

CHAFE, W.L. 1980a. ***"The Development of Consciousness in the Production of Narrative"*** *In W.L. Chafe (ed.), 1980.*

CROOKES, G. 1990. ***The Utterance and Other Basic Units for Second Language Discourse Analysis."*** *Applied Linguistics 11/2.*

CRYSTAL, D., P. Fletcher & M. Garman. 1989. ***Grammatical Analysis of Language Disability.*** *London: Whurr Publishers Ltd.*

DECHERT, H. 1983. ***"How a Story is Done in a Second Language"*** *In C. Faerch & G. Kasper (eds). 1983*

FAERCH, C., & G. Kasper (eds.). 1983. ***Strategies in Interlanguage Communication.*** *Harlow: Longman.*

FULCHER, G. 1987. ***"Tests of Oral Performance:*** *The Need for Data-based Criteria". ELT Journal, 41/4.*

GOLDMANN-EISLER, F. 1961. "Hesitation and Information in Speech" In C. Cherry (ed.) Information Theory. Butterworth: London.

LENNON, P. 1990. "Fluency and the Advanced Learner" Language Learning. 40/3.

LINDE, C., & W. Labov. 1975. "Spatial Networks as a Site for the Study of Language and Thought". Language, 51/4.

LONG, M.H. 1989. "Task, Group and Task-group Interactions". University of Hawai'i Working Papers in ESL. 8/2.

LONG, M.H., & P.A. Porter. 1985. "Group Work, Interlanguage Talk and Second Language Acquisition" TESOL Quarterly, 19/2.

NUNAN, D. 1989. Designing Tasks for the Communicative Classroom. Cambridge: CUP.

OCHS, E. 1979. Planned and Unplanned Discourse In T. Givon (ed.) Discourse and Syntax. New York: Academic Press.

POULISSE, N., & E. Schils. 1989. "The Influence of Task and Proficiency-Related Factors on the Use of Compensatory Strategies: A Quantitative Analysis." Language Learning 39/1.

SNOW, C. 1987. "Beyond Conversation" In J.P. Lantolf & A. Labarca (eds.) Research in Second Language Learning: Focus on the Classroom. Delaware Symposium 6. Norwood, N.J.: Ablex.

TANNEN, D. 1980. "A Comparative Analysis of Oral Narrative Strategies: Athenian Greek and American English" In W.L. Chafe (ed.), 1980.

WEIR, C., & M. Bygate. 1990. "Meeting the Criteria of Communicativeness in a Spoken Language Test" Workshop paper presented at RELC Conference, 1990.

THE LEARNER'S EFFORT IN THE LANGUAGE CLASSROOM

N S Prabhu

It is normal in the discussion of language pedagogy to focus attention either on teaching activities or on learning theories. In discussing teaching activities, we are concerned with what the teacher does or gets done by learners in the classroom and, at a remove, with what supports, demands or constrains the teacher's work, in the form of syllabuses and materials. In the discussion of learning theories, we are concerned with what might happen to the learner's language ability as a result of such teaching activities or, more widely, as a result of encounter with the language anywhere. These two areas of pedagogic discussion can be said to constitute two dimensions of the notion of a 'method': a method is, on one dimension, a set of classroom procedures for the teacher to carry out and, on another dimension, a concept or theory of language learning which informs or justifies those teaching procedures. Accordingly, we classify different views or formulations of language pedagogy as different methods, i.e. different combinations of teaching procedure and learning theory.

There is, however, one other aspect of language pedagogy which can usefully be brought under focus and, when examined, can reveal a meaningful basis for linking teaching activities and learning theories. This has to do with the nature of the learner's effort in carrying out classroom activities - not with the teacher's effort or intention in setting the activities, nor with the hypothesised psychological impact of those activities, but simply with what the learner has to do to perform them successfully. The form of an activity set by the teacher does not always indicate correctly the nature of the effort demanded of the learner in its performance, and the teacher's intentions about the learner's effort may often be at variance with the actual effort the learner has to make. Any impact that a teaching activity has on the learner's language development, however, has to be the result of what the learner does, rather than of what the teacher does or intends the learner to do. An examination of the learner's effort is thus of some importance in relating teaching activities to learning theories.

The purpose of this paper is to identify different forms of effort by the learner in the language classroom, and to suggest what may be called a broad typology of

learner effort. I will also try to point out the possible benefits of bringing such a typology to bear on the discussion of teaching activities and learning theories.

Imitation versus Cerebration.

Let me begin with a very general distinction between two types of effort by the learner, which I will call 'imitation' and 'cerebration'. Imitation is a matter of the learner attempting to do, as accurately as possible, what the teacher does in an exemplary role - such as producing a sound or pattern of sounds as exemplified, repeating words or sentences, altering sequences of words as demonstrated, copying strings of words or sentences from the textbook or blackboard into a note book, or conforming to a given format of writing (such as the format of a letter or job application). The effort is to achieve a close resemblance to the model provided - to conform to it as completely as possible. The language learner imitates the behaviour of the language knower in order to become a language knower himself/herself, in due course. Cerebration, in contrast, involves the learner thinking things out - finding ways of distinguishing between sounds or sound sequences, looking for workable rules for putting together word sequences, puzzling out reasons why sentences are right or wrong, applying given rules to make choices between alternatives, guessing the meaning of a word, making the most sense of what is heard or read, performing a piece of reasoning on the strength of what is known or can be inferred, or putting meaning across as well as can be done with available resources of language. Cerebration is essentially effort to cope with problems - to make sense, to work things out, and to put solutions to a test. The language learner attempts to become a language knower, not by copying the latter's language behaviour, but by constructing the knowledge-system that underlies that behaviour.

There have, as we know, been pedagogic approaches which aimed primarily, or exclusively, at the effort of imitation by the learner. The audio-lingual method was, at one stage, more frankly labelled as the mimicry-memorization ('mim-mem') method,and the term 'language drill' clearly refers to an activity in which the learner is called on to imitate repeatedly and quickly. Dialogues have been used to get learners to imitate longer stretches of language, and there are modernised versions of the same basic activity today, in the form of role-play activities. There have also been pedagogic activities aiming mainly or entirely at the effort of cerebration by the learner. The grammar-translation method of the past century, and the classical-scholastic procedures that preceded them, both involved getting

learners to puzzle out parts of the language system, as did later teaching procedures which were labelled 'cognitive-code'. These approaches too have their modernised versions, in the form of 'grammatical consciousness-raising', 'focused learning', 'instructed language acquisition' and task-based teaching. Indeed, one way of understanding the periodical pendulum swings in the history of language pedagogy is to see the profession as moving its bets, as it were, between the effort of imitation and the effort of cerebration (or hedging its bets between the two in some way).

While imitation and cerebration can be related relatively easily to different teaching approaches, it is less straightforward to relate them to different learning theories. On the face of it, imitation implies the theory that language learning is a matter of copying the language knower - with some effort to begin with, but with less and less effort as recurrent attempts are made, leading eventually in an effortless, automatic copying of the behaviour involved. This is the familiar notion of habit-forming as a process of language learning. Similarly, cerebration seems to imply the theory that language learning is a conscious process of rule-discovery and rule-application, the rules discovered at different stages eventually forming a less conscious knowledge-system able to support language behaviour. This is the notion of learning a language by internalizing its grammar. However, this is only an obvious interpretation. It is equally possible to see imitation as a facilitator of internal grammar construction: repeated copying of a piece of language, it can be argued, brings about a subconscious abstraction of its structural features, leading to rule-formation. The fact that the learner's effort consists of imitation need not mean that language-learning itself, or later language-use, is a matter of mere imitation. Indeed, if one looks at the early ideas of language drills and habit-formation, developed by such thinkers as Harold Palmer in the early part of this century, one can see the notion that a 'habit' is in fact a principle of sentence-construction, or pattern, abstracted subconsciously from the set of sentences imitated, and employed thereafter to construct new sentences (Palmer, 1921). Imitation, therefore, may well be thought of as a favourable condition for internal system-development, far from being a denial of it. Similarly, cerebration need not imply the theory that language-learning is a process of conscious grammar-construction which then becomes a subconscious knowledge system. There can be forms of cerebration which do not focus at all on the language as such; and even forms of cerebration which do focus on language can be viewed as merely favourable circumstances, or useful triggers, to a process of subconscious grammar-development, as done in arguments for grammatical conscious-raising (Rutherford, 1987).

There is, in passing, one area of language pedagogy where imitation points clearly to the theory that learning is a matter of copying behaviour. I am referring to the teaching of language for specific purposes, that is to say, the teaching of specific, and limited, language for specific, and limited, use. If limiting the learner to specific forms, samples or uses of language is seen as creating favourable circumstances for internal system development, there would in effect be a claim about the possibility of limited internal systems, delimitable by pedagogic decision - a claim larger than has been made in the discussion of specific-purpose language teaching. Therefore, teaching language for specific purposes must mean getting the learner to copy the typical behaviour of language-users in specific situations, with imitation as the most relevant learner-effort. The teaching of limited language thus implies what may be called a limited theory of language learning - one which regards copying as the learning process.

Reproduction versus Simulation.

Let me now go on to a further distinction in learner effort, within the category which I have called imitation. I would like to distinguish between imitation of the language form itself, which I will call 'reproduction', and imitation of language behaviour with attention to its context, which I will call 'simulation'. Reproduction thus aims at linguistic accuracy or approximation in the act the copying, while simulation aims at a match between language form and its context, i.e. at appropriacy. The kinds of activity that are normally called language-practice, or perhaps uncontextualised language practice - such as a repetition drill, the use of a substitution table, memorization and recall of language samples, sentence construction on a given model, some forms of heavily guided composition (which make copying look like composing) - are all a matter of reproduction, in the way I am using the term. Many kinds of activity which are referred to as communicative practice, or 'structural practice with functional honesty', to use Johnson's (1982: 109) term - such as role-playing a given dialogue, making choices between given expressions in the light of such requirements as politeness and formality, writing a letter, report or job application by imitating a given model - are all a matter of simulation. In reproduction, learners are occupied with linguistic forms; in simulation, they are occupied with what may be regarded as an indexing of linguistic forms with their communicative values or functions. Reproduction can be said to lead to the ability, as Newmark (1966) saw it, to ask a stranger for a light with "do you have a light?", "Do you have fire?" or "Are you a match's owner?",

while simulation is meant to lead to a choice of "Do you have a light?" or "Got a match?" in preference to the other expressions.

Reproduction has clearly been a central part of the Structural Approach to language pedagogy, though it has had a stricter adherence in the Audiolingual Method, arising from the work of Fries and Lado, than in the Situational Method advocated by Hornby, which represented some attempt at simulation. Arguments for a central place for simulation have been advanced, and accepted fairly widely, in the last 20 years, under the label Communicative Approach. Following normal rhetorical tactics, the Communicative Approach has been presented as a challenge to the Structural Approach and as a radical replacement for it. It is therefore important to realise that replacing the Structural Approach with the Communicative Approach is, in terms of learner effort, replacing reproduction with simulation. The general nature of the learner's effort is still imitation - a matter of copying language behaviour. Further, although simulation represents, from one point of view, a widening of pedagogic intentions based on a less restricted view of language ability, it also represents, from another point of view, a more restricted view of language and a more limited theory of learning. Learners' imitation of language forms can, as we have noted, be seen as a favourable condition for internal system development, the system so developed forming the basis of subsequent language behaviour. Reproduction, that is to say, permits the view that language learning is a matter of internalizing a system of rules, not just a matter of copying the language knower's behaviour. It is difficult to see how such a view of learning can be reconciled with simulation, since the situational appropriacy of language forms is at best a matter of established convention in a language community, falling far short of being a productive system. When learners are able to say to a stranger, "Do you have fire?", "Have you got illumination?", "Are you a match's owner?" etc., they are giving evidence of an internalized grammatical system, since they are producing well-formed linguistic expressions which they could not have merely copied from a language knower. When, on the other hand, learners are able to choose the situationally appropriate expressions "Do you have a light?" or "Got a match?", they can only be copying a language knower's behaviour in a similar situation, simply because it is impossible to see what system of rules could lead to a choice of appropriate expressions over inappropriate ones, without any copying of behaviour as such. One can perhaps point to the possibility of a rule system of pragmatics underlying some aspects of appropriacy in language behaviour, but it is still fair to say that a very large part of situational appropriacy is a matter of ad hoc convention, rather than rule-governed behaviour. In any case, pedagogic arguments in support of simulation have consistently appealed to conventional appropriacy, rather than to

any rule system underlying appropriacy - indeed, they have laid stress on the fact that appropriacy is not a matter of rule systems and therefore needs to be acquired through simulation. If this is so, the claims that can be made for simulation in preference to reproduction are not very large. Simulation is preferable to reproduction only in terms of pedagogic intentions. In terms of learner effort, it is as much a matter of imitation as reproduction; and, in terms of learning theory, it is much more limited than reproduction, since it regards learning itself as no more than a copying of behaviour.

Construction versus Deployment.

I now wish to make one further distinction in learner effort, this time within the category of cerebration. We can, once again, distinguish between cerebration focussed on language itself, which I will call 'construction', and cerebration focused on a knowledge of the world handled through language, which I will call 'deployment'. 'Construction' refers both to an effort by the learner to understand some part of the language system - a matter of construing the system - and to any effort to put together linguistic expressions on the strength of that understanding. The term 'exercise', in contrast to the term 'practice', generally refers to the effort of construction - an activity in which the learner is challenged to understand and demonstrate that understanding. An exercise is like a test, in that it presents the learner with risks of going wrong, unlike practice which aims to avoid or minimise such risks. All grammatical exercises, in this sense of the term 'exercise', call for the learner effort of construction, as do various forms of vocabulary or pronunciation exercises, and indeed exercises calling for an assessment or explanation of appropriacy, though these are rarely employed. In contrast, what I am calling 'deployment' involves learners handling information and developing knowledge about the world, by drawing on what information and knowledge of the world they already have, what cognitive abilities of inferring, reasoning, relating, etc, they possess, and what linguistic resources they command. Reading, in the sense of interpreting texts, involves the effort of deployment, as does free composition, in the sense of learners finding ways of saying what they have to say. The 'immersion' classroom - learners studying a curricular subject in the medium of the target language - calls for the effort of deployment, in that it makes learners acquire new knowledge through the language. Task-based activity, in the sense the term was used on the Bangalore Project (Prabhu, 1987) is a means of getting beginner-level learners to make the effort of deployment in the language classroom. Deployment is typically the effort called for in problem-solving activities. Reading,

writing, acquiring new knowledge can all be regarded as processes of problem-solving in a very general sense; task-based teaching can then be seen as a more deliberate, more focused and more manipulated form of problem-solving, as a language teaching device. It is true that what I have called the effort of construction is also a form of focused problem-solving, except that it involves solving problems of language as such, unlike deployment which involves the use of language as a resource in solving problems of world knowledge.

I pointed out earlier that cerebration in regard to the language being learnt - what we are now calling construction - has been a part of language pedagogy for a long time, in such approaches as the Grammar-Translation Method and the Cognitive-Code method. I also pointed out how the language teaching profession has, at various times, upheld or denied the value of such effort of construction for the learning of a language. It is perhaps worth asking, for a moment, what follows when the effort of construction is rejected as a pedagogic procedure. One may, in rejecting construction, reject all cerebration, as happened more than once when grammar-teaching was rejected as a means of language teaching. Such rejection of all cerebration leaves one with only imitation as a desirable learner effort, whether in the form of reproduction or simulation. Alternatively, one may deny the value of construction but still see value to cerebration. This leads to the adoption of deployment as a desirable form of learner effort. To put it differently, the learner grappling with the rules of grammar is making a cerebral effort focused on language. If, in rejecting this, one rejects both the cerebral effort and the focus on language, one is only left with relatively mindless activity, with the focus at best on a faithful copying of behaviour. If, on the other hand, one rejects only the focus on language but not cerebral effort, one has the option of setting up activity involving cerebration on things other than language - that is to say, deployment.

The effort of construction can, as we noted earlier, be related to one of two theories of learning: it can be seen either to lead the learner to a conscious knowledge of the linguistic system, which then gradually descends to a sub-conscious level, or merely to serve to trigger a separate, subconscious process of system development. In contrast to this, the effort of deployment is relatable only to the latter theory - that a subconscious process of system development is triggered and promoted when language is employed as a resource in an effort to understand the world.

Four types of learner effort.

I have now identified four types of learner effort in the language classroom - reproduction, simulation, construction and deployment. Reproduction and simulation both involve imitation - of the forms of language in one case, and of the social role of the language-user in the other. We can perhaps call classroom activities involving these two types of effort 'form-focused' and 'role-focused', respectively. The other two types of effort - construction and deployment - involve cerebration, cerebration on the rules of language in one case, and on knowledge of the world in the other. Activities involving these two forms of effort can therefore be called rule-focused and meaning-focused, respectively. From a different point of view, we can group together the effort of reproduction and that of construction, both being focused on language as an autonomous phenomenon. They differ between them in that one involves a copying of the phenomenon while the other involves trying to make sense of it. Terms such as 'grammar-teaching' or 'the teaching of structure' are often used ambivalently - to refer to reproduction and/or construction - thus obscuring a major difference in terms of learner effort, namely, that between imitation and cerebration. Similarly, simulation and deployment can be grouped together, as forms of effort focused on things outside language as a formal entity. The focus, in one case, is on the performance of a social role, and in the other on an understanding of the world. The term 'communicative' is often used ambivalently to refer to either or both of them, thus blurring the distinction, once again, between imitation and cerebration.

Usefulness of the Typology.

What is the usefulness of this categorization of learner effort? First, it enables us to see known forms of language pedagogy in a somewhat new light, or from a newer point of view. Looking at things from as many viewpoints as possible is one of the mental tools we have for increasing understanding. It is, for instance, common in the discussion of language pedagogy to contrast a formal, grammar-based approach with an apparently more comprehensive communicative approach directed to real-life language use. But this is a contrast based on teaching intentions, nor on learning theories. An examination of learner effort leads us to a different contrast, that between imitation and cerebration, which is likely to be more significant for the process of learning. We then begin to see that both a grammar-based approach and a communicative approach can involve no more than a copying of behaviour and that, between the two approaches it is the grammar-based

approach which permits a non-behaviourist theory of learning even when it involves only the effort of imitation.

Secondly, a categorization of learner effort provides us with an additional frame of reference to which we can usefully relate concepts and issues in pedagogy. Take, for instance, the concept of authenticity. The prevalent notion is that authenticity is a characteristic of the samples of language presented to the learner: the samples are authentic if they are taken from actual instances of language use. This notion has been challenged by Widdowson (1978: 79-80) who regards authenticity, instead, as a characteristic of the learner's engagement with language samples: the learner's engagement is authentic if the learner responds, interacts or copes with language samples in the same way that language users do. Given our categories of learner effort, we can see that the first notion of authenticity makes sense when the learner's effort consists of simulation, while Widdowson's notion clearly envisages the effort of deployment. What is involved therefore is not just a terminological difference but a difference in theories of learning: simulation-based pedagogy cannot possibly adopt Widdowson's notion of authenticity simply because simulation cannot be claimed to be the central characteristic of real-life language use. Or take Widdowson's well-known distinction between 'usage' or 'use'. The distinction is clearly between language-focused and non-language focused learner effort - between reproduction and construction on the one hand, and simulation and deployment on the other. However, the distinction leaves out the equally important contrast between imitation and cerebration, though Widdowson elsewhere takes a stand against simulation-based ESP, and in defence of grammatical problem-solving. The usage-use distinction, therefore, is only a distinction between reproduction and deployment, though it is presented as a more global one. Similarly, Krashen's (1981) distinction between 'learning' and 'acquisition' is only a distinction between construction and deployment, in our terms; it ignores reproduction and simulation, with the result that its proponent often finds himself recommending simulation procedures as if they were deployment procedures. We can also relate the issue of error-treatment in language pedagogy to our categorization of learner effort. Imitation-based pedagogy (reproduction and simulation) is necessarily committed to error-prevention: if language learning takes place through an imitation of language behaviour, then it is crucial for the imitation to be accurate; any errors in imitation can, in principle, make the learner an erroneous language user. Cerebration-based pedagogy, on the other hand, has necessarily to permit errors by learners: problem-solving would be a mere pretence if it did not permit trial and error. Further, construction-based pedagogy has a commitment to dealing with errors when they occur - e.g. setting

further exercises or promoting a better understanding of the rules, in some way - while deployment-based pedagogy is necessarily error-tolerant, since dealing with errors as such will undermine the genuineness of deployment.

Finally, let me point out briefly a possible use of this typology of learner effort in the preparation of teaching materials. It is common to find, in textbooks for language teaching, an apparently varied pattern of activities for the learner, under such labels as comprehension, vocabulary, grammar or structure, pronunciation, and composition. In the more modernised textbooks, one also finds sections on such communicative activities as role-play, writing an office report, etc. A closer examination of such materials, however, often reveals that most of the comprehension questions demand of the learner little more than a copying of relevant parts of the reading text (a question merely serving as a clue to which part of the text is to be copied), the vocabulary exercise calls mainly for a copying of words or an imitation of given sentences by incorporating given words in them, that the grammar or structure work involves a piece of intensive 'practice' (that is to say, reproduction in some form), that the pronunciation work is understandably a matter of accurate imitation, that the communicative activity of role-play is a matter of learners (working in pairs in keeping with enlightened pedagogy) mouthing to each other their parts in a scripted dialogue, and that the writing of an office report is a matter of imitating faithfully all the essential parts of a model office report, making only some substitutions of content words to alter the information content. It is clearly the textbook writer's intention to employ a variety of teaching activities (and indeed, different teaching approaches - both linguistic and communicative), but, from the viewpoint of learner effort, what all or most of the activities call for is just imitation. Perhaps it would be useful for textbook writers to bear a typology of learner effort in mind and to ensure that the activities provided do involve the type of effort which they see the most value to - or, if they wished to be eclectic, to ensure that there are some activities which involve each type of learner effort. A typology of learner effort can thus serve both as a further dimension to pedagogic understanding and as a further aid to pedagogic practice.

REFERENCES

*JOHNSON, Keith, 1982. **Communicative Syllabus Design and Methodology**, Oxford: Pergamon.*

*KRASHEN, Stephen D, 1981: **Second Language Acquisition and Second Language Learning**, Oxford: Pergamon.*

*NEWMARK, Leonard, 1966: **"How not to interfere with language learning"**, International Journal of American Linguistics, 32, I,II.*

*PALMER, H.E., 1921: **The Principles of Language Study**, (Reprinted) Oxford: Oxford University Press, 1964.*

*PRABHU, N.S., 1987: **Second Language Pedagogy.**, Oxford: Oxford University Press.*

*RUTHERFORD, William E, 1987: **Second Language Grammar: Learning and Teaching**, London: Longman.*

*WIDDOWSON, H.G., 1978: **Teaching Language as Communication**, Oxford: Oxford University Press.*

Diary Studies of Classroom Language Learning: The Doubting Game and the Believing Game[1]

Kathleen M. Bailey

Introduction

Recently, after giving a talk about language classroom research at a university in Japan, I was asked during the question and answer period, "Why have you forsaken the diary studies?" I was somewhat taken aback by the underlying proposition in the question: Although I hadn't published anything on the diary studies for some time, I had no awareness of having "forsaken" them.

In contrast, I have recently been introduced at several conferences as being well known for my work on the diary studies. This too surprised me, but a colleague pointed out that the language learning diary studies were my "brainchild." My response was that the Schumanns had conceived of the genre, but they left it on my doorstep.

My intent in writing this paper is to clarify where I stand and what I believe about the language learning diary studies. In doing so, I hope to highlight issues of interest for second language acquisition researchers, for language teachers, and for language learners.

What is a diary study? According to Bailey and Ochsner,

> "A diary study in second language learning, acquisition, or teaching is an account of a second language experience as recorded in a first-person journal. The diarist may be a language teacher or a language learner -- but the central characteristic of the diary studies is that they are introspective: The diarist studies his own teaching or learning. Thus he can report on affective factors, language learning strategies, and his own perceptions -- facets of the language learning experience which are normally hidden or largely inaccessible to an external observer (1983:189)."

The diary studies are thus first-person case studies -- a research genre defined by the data collection procedures: A language learner keeps an intensive journal using

introspection and/or retrospection, as well as observation, typically over a period of time. The data analysis may be done by the diarist himself or by an independent researcher using the learner's diary (or some "public" version of that diary) as data.

As noted above, language diaries can be kept by teachers as well as learners. In this paper, however, I will not be concerned with teachers' diaries. For information on that topic the reader is referred to Bailey (1990), Enright (1981), or Nunan (1989:55-60). I will, however review the work of some teachers who kept journals of their experiences as language learners and who have commented specifically on how that experience related to their own teaching (e.g., see Danielson, 1981).

Neither will this paper address the use of dialogue journals as pedagogic tools. In dialogue journals, teachers respond to what students write in their diaries, as part of an on-going exchange. For further information on this procedure, see Kreeft-Peyton (1990), Kreeft-Peyton and Reed (1990), Popkin (1985), Spack and Sadow (1983), and Staton (1981).

Language Learning Diary Studies

When learners keep journals of language learning experiences in formal instructional settings, the resulting diary studies fit the tradition of language classroom research as described by Allwright (1983), Bailey (1985), Allwright and Bailey (1991), Brumfit and Mitchell (1990), Chaudron (1988), Gaies (1983), Long (1980, 1983), Mitchell (1985), and van Lier (1984, 1988, 1989). The diary studies are a form of empirical (data-based) research, enhanced by introspection, in the tradition of naturalistic inquiry (Guba, 1978; Lincoln and Guba, 1985)

Naturalistic inquiry is a broad rubric which can cover many different sorts of investigation. In classroom research these include ethnography, ethnomethodology, some discourse analyses and some case studies. In naturalistic inquiry, "first no manipulation on the part of the inquirer is implied, and, second, the inquirer imposes no a priori units on the outcome" (Lincoln and Guba, 1985:8). In other words, (1) people are studied in naturally occurring settings, rather than in randomly sampled groups created artificially for the purposes of an experiment, and (2) the analytic categories typically emerge from examining the data. Allwright and Bailey (1991:40-45) have compared naturalistic inquiry and experimental research as follows: In the latter, researchers exercise a high degree of control over variables

and (in the true experimental designs) exert an intervention, while in the former, researchers choose not to control variables and try not to intervene. A similar contrast is drawn by van Lier (1988:56-60; 1989:33-35), who compares the naturalistic and experimental approaches to research on two intersecting continua: selectivity (from structured to unstructured) and intervention (from controlled to uncontrolled).

In a discussion of the methodological basis of introspective methods, Grotjahn (1987) locates introspective research among other forms of research by using categories from Patton (1980). In this framework, a research approach can be categorized according to:

(1) its design (non-experimental or "exploratory," pre-experimental, quasi-experimental, or true experimental);

(2) the type of data used (qualitative or quantitative);

(3) the sorts of analyses involved (interpretive or statistical).

The language learning diary studies can be classified as being in Grotjahn's (1987:59) "exploratory-interpretive" category, in that they typically utilize non-experimental designs, qualitative data, and interpretive analyses. A few that have utilized quantitative analyses are considered "exploratory-quantitative-statistical" in nature (e.g., Brown, 1983, 1985a; Matsumoto, 1989; Parkinson and Howell-Richardson, 1990). Others (e.g., Ellis, 1989a; Matsumoto, 1989; Schmidt and Frota, 1986) have involved combinations of data types and both qualitative and quantitative analyses.

I am using the term "diary study" in a way different from Hatch's use of the phrase in introducing the language acquisition case studies reprinted in Hatch, 1978. In those instances, researchers (often parents) kept daily observational logs of young learners' development. Although I will use the terms diary and journal interchangeably, here diary study will be restricted to the situation in which the learner himself keeps the intensive journal, thus permitting introspection and retrospection to inform the process of observational data recording.

It is important to note that a learner's diary alone doesn't constitute a diary study. The diary is typically only the data. In order to be considered a diary study, a paper must include an analysis. Rivers published her diary of a six-week trip in

Latin America, documenting her experiences "learning a sixth language" (Rivers, 1979, reprinted in 1983). Fields (1978) published a diary of her experiences in a Berlitz Spanish class and on a subsequent trip to Mexico. These papers include no explicit analysis of the data, so while the journal entries are interesting and certainly available to be analyzed, I do not consider these articles to be actual diary studies. Of course, in some instances the student may analyze the learning experiences within the journal itself. This is the nature of some of the diary entries made by Fields (1978).

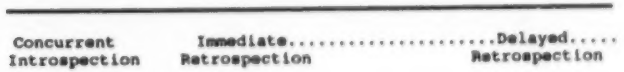
The diary data are a combination of learners' records of events and their interpretations of those events. The learners' introspection permits the reader to understand some aspects of language learning which are normally hidden from view. Introspective methods encompass "self-report, self-observation, and self-revealment" (Grotjahn, 1987:55). As Seliger has pointed out (1983:183), "introspections are conscious verbalizations of what we think we know."

A problematic methodological issue is timing. When does the verbalization take place relative to the event about which the learner is introspecting? The term introspective data is held by many researchers to refer only to data "gathered from subjects while they carry out a task" (Fry, 1988:159, underscoring added). In comparison, retrospective data are those "collected after the event" (ibid.). Of course, the difficulty here is that true introspection (such as the think-aloud protocols used by Abraham and Vann, 1987; Cavalcanti, 1982; Cohen, 1987; and others) can take place only so long as the event is occurring. Some diarists have reported making notes during a class or conversation -- e.g., Bailey (1980), and Henze in Rubin and Henze (1981) -- but there is concern that this procedure may detract from the language learning process. As Fry notes, "reporting on how one is doing a task while doing it is a double task" (ibid:160).

In contrast to true introspection concurrent with the task, the term retrospection involves a very broad data collection timespan, ranging from immediately after the event (following a language class, for example) to years later (as is the case in the language learning histories). Mann (1982:87) identifies three basic techniques for using verbal self-report as thinking aloud, introspection, and retrospection. Cohen and Hosenfeld (1981) distinguish among three similar categories of introspective data collection, each of which represents a band rather than a point: introspection (during the event), immediate retrospection (right after the event), and delayed retrospection (hours or more following the event). Thus the cover term

introspection entails all three zones: concurrent introspection, immediate retrospection and delayed retrospection, as depicted in Figure 1:

Figure 1: Introspection Immediacy Continuum



Data from this entire immediacy continuum have been utilized in the diary studies.

Given this general background on language learning diary studies, we can now turn to the purpose of this paper. It is my intent here first to review some of the language learning diary studies published to date, and then to examine the "pros and cons" of the diary studies by adopting first a critical attitude and then a more accepting stance.

The Doubting Game and the Believing Game

The title of this paper mentions "the doubting game and the believing game." These concepts were introduced by Elbow (1973), in the appendix to his book, Writing Without Teachers. Elbow's ideas originally referred to literary criticism but Elbow felt they could be applied equally well to "most procedures in the humanities and social sciences" (ibid:166). I find the doubting and believing game images useful in characterizing possible attitudes toward incoming information from research as well. A decade ago, Larsen-Freeman (1981) used these game metaphors in her comparison of four prominent theories of second language acquisition (SLA). With each theory she first adopted a critical stance, doubting and questioning the author's position (the doubting game). She then asked what insights she could gain about second language acquisition by adopting the author's position uncritically (the believing game).

The doubting game "seeks truth by indirection -- by seeking errors" (Elbow, 1973:148). This position underlies the logic of the null hypothesis in experimental research. (The formulaic language of the null hypothesis typically begins with the expression, "there will be no statistically significant difference between..." or "there will be no statistically significant correlation between...".) In this tradition, the researcher's job is to conduct an analysis which attempts to reject the null

hypothesis. The doubting game also emphasizes objectivity - the separation of self (the subject) from the object under investigation. It takes a critical, questioning stance toward evidence and conclusions drawn from the data.

In contrast, the believing game "emphasizes a model of knowing as an act of constructing, an act of investment, an act of involvement" (Elbow, 1973:173). It seeks truth by affirmation. In this sense, truth in the believing game is related to myth, metaphor, and allegory -- to convincing forms of story (Reason and Hawkins, 1988). In taking this attitude, the listener/reader accepts the experience of the person making the assertion. The believing game is "not an act of self-extrication but of self-insertion, self-involvement -- an act of projection" (Elbow, 1973:149). Playing the believing game involves inclusion of the subject as a legitimate focus of investigation. In fact, the believing game "is built on the idea that the self cannot be removed: Complete objectivity is impossible" (ibid.:172).

The objectivity/subjectivity continuum is crucial in interpreting the diary studies. In this approach to research, the subject becomes the object: We conduct (and read) diary studies to understand language learning as seen by the learners. We will return to this problem when we apply the metaphors of the doubting game and the believing game to the diary studies, after reviewing the findings of several such studies.

A Review of the Language Learning Diary Studies

This literature review will be limited to those language learning diaries which have been published. There are numerous unpublished manuscripts which employ the diary study method, but they are typically very long, unwieldy documents and their accessibility is limited. For these reasons, I will review only those diary studies that I have located which are in print and are therefore readily available to the academic reading public.[2]

The language learning diary studies can be divided into two groups: (1) those in which the diarist and the analyst are the same person, and (2) those in which the researcher analyzes journals kept by other language learners. Matsumoto (1987) has called these "introspective" and "non-introspective" diary studies, respectively. I am concerned that Matsumoto's labels may lead to some confusion, but she is careful to point out that in her usage, the terms introspective and non-introspective refer only to the data analysis, and not to the data collection phase of the research.

The terms direct analysis and indirect analysis have been suggested by van Lier (personal communication, 1991).

Diary Studies with Introspective (First-Person) Analyses

Pioneering work with the diary studies was published by Schumann and Schumann (1977), who kept intensive journals of their experiences in three language learning contexts: studying Farsi in Los Angeles, and in Iran, and studying Arabic in Tunisia. In this first analysis, the researcher/diarists identified six personal variables of importance in their language learning: the role of materials, rejection of the teaching method, and nesting patterns (F. Schumann); and a preference for eavesdropping vs. speaking, the desire to maintain a personal language learning agenda, and transition anxiety (J. Schumann).

What is noteworthy about this brief early report is the striking difference between two learners undergoing essentially the same language learning experiences. For example, F. Schumann found she could not begin to cope with the target language in a new environment (e.g., Tunisia and Iran) until her physical surroundings were ordered and comfortable (a phenomenon she referred to as her "nesting patterns"). In contrast, J. Schumann coped with his "transition anxiety" by plunging into the study of the target language immediately, regardless of his surroundings.

In 1980, F. Schumann published a further analysis of her original journal data. Additional themes which emerged as important in her language learning were (1) the role of the expatriate (English-speaking) community; (2) difficulties of a woman language learner in obtaining input, given social taboos against interaction; (3) the difficulties of obtaining input as a native speaker of English since most potential interlocutors wanted to speak English instead of the learner's target language; and (4) cooperation vs. competition in language learning. This last theme was later investigated in other learners' diaries by Bailey (1983; see below).

My first venture into the diary studies (Bailey, 1980) occurred when I was a student of J. Schumann. At that time I took a thirty-hour reading course in French as a foreign language to prepare for an examination. I also kept an intensive journal of this class, which could be characterized as a lower-intermediate French course. Factors that emerged as important in my language learning in that context were (1) the language learning environment (both physical and social), (2) my preference for

a democratic teaching style, and (3) the importance of success and positive feedback in the second language learning process.

As I reread that paper now, over a decade after writing it, I wish I had included more excerpts from the journal. What I find most compelling at this point is the account of a classroom crisis, taken directly from the diary, and the discussion of how the resolution of that crisis influenced the rest of the course. The argument (which occurred when the teacher returned a test that the students considered unfair) provides an example of what is known as a "natural experiment" in ethnography. As Hammersley and Atkinson explain, a natural experiment is an opportunity to explore "some unusual occurrence" (1983:31). These unusual but naturally occurring events

> "reveal what happens when the limiting factors that normally constrain a particular element of social life are breached. At such times social phenomena that are otherwise taken-for-granted become visibly problematic for the participants themselves, and thus for the observer" (ibid.:32).

The diary entry about the crisis (Bailey, 1980:60-61) documents the French students' verbal assertions that the test was unfair and too long; that tests are devastating to learners if the teacher's intent is to show them how little they know; that the teacher thought we weren't "very bright" and that the class was becoming "an armed camp." The natural experiment, in this case, allowed me to observe a group of very angry students -- hardly an event one could precipitate in an experimental treatment (provided that one was concerned about being an ethical researcher).

In 1980, Jones (another student of J. Schumann) reported on her experiences as an intermediate learner of Indonesian as a second language in an intensive program. Her diary study focuses on social and psychological factors which influenced her language learning. The positive experiences of interacting with her host family stand in stark contrast to the difficulties and frustrations she faced in the formal instructional program she attended. This brief report is based on the author's masters thesis (Jones, 1977), which provides more information about the language learning context than the paper does.

Danielson reports on her experiences as an older learner studying Italian as a foreign language. She enrolled in classes at two different levels. In the more

advanced class, Danielson was challenged and learned new material. In the lower level she gained confidence and practice opportunities. Danielson was a very experienced language teacher herself when she conducted this diary study. In the final paragraph (1981:16) she comments, "The observations I have included here are admittedly quite commonplace yet they all came as a revelation to me."

Another diary study is based on the experiences of a teacher-in-training (Henze), who enrolled in an Arabic class at the same time she was a graduate student investigating language learning strategies with her professor (Rubin). The resulting report (Rubin and Henze, 1981) benefits from two points of view in the analysis of the journal entries. In several other diary studies, language learners had recorded anything they considered interesting or important. Rubin and Henze modified this procedure and conducted what they called a "directed diary study": Henze's journal observations were directed specifically toward the role of inductive and deductive reasoning in her Arabic learning.[3]

Henze focussed her journal entries by using a list of six inductive and ten deductive reasoning strategies previously identified by Rubin (ibid.:17 and 18). She used the list "as a guide but the examples were only to be taken as suggestive of the kind of detail and strategies desired" (ibid.:19). The authors report that Henze found frequent use of the list distracting, but that keeping the basic strategies in mind "was very helpful in focussing her attention" (ibid.). Although she used the inductive reasoning strategies very little at the beginning of the Arabic course, after about two months, in response to different kinds of activities (e.g., dialogues) introduced by the teacher, more examples of inductive reasoning occurred.

Schmidt and Frota (1986) also conducted a diary study. The resulting article was written by a linguist-turned-language-learner and another linguist/analyst. This paper provides an excellent example of what the diary studies can offer. It documents Schmidt's learning of Portuguese for five months in Brazil. This was one of the first published diary studies to combine journal entries with other data sets. In addition to "R's" diary, the researchers tape-recorded and analyzed periodic target language conversations. Frota, a native speaker of Portuguese, conducted an error analysis of the conversational data.

The diary itself documents three stages in R's acquisition of Portuguese. During the first three weeks in Brazil he had no instruction in Portuguese. For the next five weeks he had both instruction and interaction in Portuguese, and for the last fourteen weeks he had interaction but no instruction. Schmidt provides

background information for interpreting his experiences in Brazil by giving the reader his language learning history.

This article may seem overlong to readers accustomed to the economic reporting of statistical publications in professional journals, but this length is a natural result of the diary study method. The paper is filled with rich examples from Schmidt's journal, written in the casual, self-as-audience style of the diary. The journal entries are complemented by transcribed conversations with Frota, and with numerical analyses of those data.

Another paper which may be classified with the language learning diary studies was written by Moore (1977). The author was a clinical psychologist and a native speaker of English, who took a post at a university in Denmark. This article discusses how Moore's proficiency in Danish influenced his professional life. The data he shares are from notes he made while attending a Danish class and after interactions with Danish colleagues: He does not specifically mention keeping a daily diary. Nevertheless, the article is informed by both introspection and retrospection, and gives the reader a sense of Moore's empathy as well as his psychology. He writes about how his experiences allowed him a better understanding of the problems faced by deaf people, aphasics, immigrants, and "especially perhaps the child in a class where the work is too difficult for him" (1977:107).

As mentioned earlier, Rivers (1979, reprinted in 1983) kept a diary of her experiences learning Spanish as a sixth language during a trip to South America. The article presents her daily diary, just as it was written, without any subsequent analysis. The entries consist of lists of the author's observations about her own strategies and hypotheses as she interacted with Spanish-speaking people. Some entries are extremely short. For example, one night Rivers wrote, "I felt I could understand the advertisements on the radio tonight" (1983:176). As a reader, I find the brevity of this entry frustrating, but parts of the journal have more depth.[4]

The second such journal (without an explicit analysis) was published by Fields in 1978 as a continued article in two issues of the Chronicle of Higher Education. The two-part story consists of Fields' diary entries as she took a Spanish course to prepare for an assignment as a newspaper reporter in Mexico, and of her trip to Mexico. While the paper lacks the detailed analysis to be properly called a diary study (as does Rivers, 1979, 1983), Fields' journal entries provide fascinating and

candid commentary about an adult learner's experiences in a formal instructional setting and then in the actual target culture.

In all of the papers discussed above, the analysis of the journal entries was done by the same person who kept the diary. The diary studies by Schmidt and Frota (1986) and Rubin and Henze (1981) each involved one learner who was also one of the analysts. The papers by Rivers (1979, 1983) and Fields (1978) both lack a formal analysis. In summary then, Table 1 lists those diary studies which have been reviewed thus far:

Table 1

Diary Studies by Learner/Diarist/Analysts

AUTHOR(S)/DATE(S)	LEARNER(S)	TARGET LANGUAGE(S)
F. Schumann & J. Schumann (1977)	2 experienced linguists/teachers	Farsi & Arabic
F. Schumann (1980)	1 experienced teacher/linguist	Farsi & Arabic
Bailey (1980)	1 experienced teacher/linguist	French
Jones (1980)	1 experienced teacher/linguist	Indonesian
Danielson (1981)	1 experienced teacher/linguist	Italian
Rubin and Henze (1981)	1 experienced teacher/linguist	Arabic
Schmidt & Frota (1986)	1 experienced teacher/linguist	Portuguese
Moore (1977)	1 psychologist	Danish
Rivers (1979, 1983)*	1 experienced teacher/linguist	Spanish
Fields (1978)*	1 reporter	Spanish

*(No explicit analysis was provided.)

Diary Studies with Non-introspective Analyses

The second macro-category of diary studies consists of those in which an external researcher analyzes the journals of other language learners. In this approach to research, which Matsumoto (1987) has called "non-introspective" diary studies, the learners' journal entries provide both the data and an "emic" (insider's)

view of language learning, while the researcher's use of SLA theory and previous research can provide an "etic" interpretation in the analysis. (For more information on the emic/etic contrast, see Watson-Gegeo, 1988:579-582, and van Lier, 1990:42-43.)

To my knowledge, the first published analysis of other learners' diaries was my work on competitiveness and anxiety in adult second language learners (Bailey, 1983). Curiously, this research started out to be an author-analyzed study on quite a different topic. I had originally planned to go back to my French class diary and document the learner's perspective on error treatment. When I analyzed the journal for references to this topic, there were very few. Although error treatment intrigued me as a researcher, it had apparently not been particularly important to me as a learner.

What I found instead were numerous comments about feeling competitive and anxious in the classroom. I was so uncomfortable with the results of this analysis that I felt compelled to look at the journals of other language learners, to see if anyone else had reported having these experiences. With their permission, I read the journals (or the public reports) written by ten other learners. The analysis led to a description of competitiveness and to two related suggestions: (1) language classroom anxiety (see Gardner, Smythe, Clement and Glicksman, 1976) "can be caused and/or aggravated by the learner's competitiveness when he sees himself as less proficient than the object of comparison" (Bailey, 1983:96), and (2) "as the learner becomes, or perceives himself as becoming, more competent (that is, better able to compete),[5] his anxiety will decrease" (ibid.). By trying to depict the relationship between competitiveness and anxiety which I found in the journals of these eleven learners, I came up with a visual model which I have since used with teachers and students, who are encouraged to trace their own "routes" through the flowchart.

Anxiety as a theme in the diary studies was also addressed by Parkinson and Howell-Richardson (1990), who used students' journals as the data base in a research project on learner variables. Their work involved two groups of English learners in Scotland: 23 students in a pilot study and 51 students in two other cohorts (29 students in Autumn, 1986, and 22 students in Spring, 1987). The authors also reported on work in progress involving the diaries of local Edinburgh people studying modern languages, including Spanish and French. The diaries of this latter group of learners were analyzed for comments on the use of the foreign language outside of class and for the learners' reported use of strategies.

The English learners' diaries were analyzed for (1) the reported use of English outside of class, (2) references to anxiety, and (3) informativity, which was a category related to specific information regarding the students' newly acquired knowledge. After quantifying the data, the researchers found a correlation between the learners' "rate of improvement and the amount of time which students spent outside class in social interaction with native speakers of English" (Parkinson and Howell-Richardson, 1990:135).

More interesting than their quantitative analyses, however, are the researchers' discussions of the data analysis procedures (ibid.:129-134) and their interpretive comments about the diary entries. For example, with regard to the variable of time spent on out-of-class activities in English, they write,

> "The figures recorded in the diaries are clearly subjective approximations of the actual length of time spent engaged in any one activity. A further variable to be taken into account is the value placed on various activities by the student himself. Clearly what a teacher or researcher may regard as 'linguistically relevant' is not always valued as such by the student diarist..." (ibid.:134).

I find the Parkinson and Howell-Richardson report tantalizingly brief and even sketchy in spots. However, the strengths of the paper include its use of quotes from the learners, the attempt to combine quantitative and qualitative analyses, and the use of data from multiple learners. The authors also conducted a pilot study to generate needed revisions in the research procedures. For more information about these projects, see Howell-Richardson and Parkinson (1988).[6]

The use of multiple learners' diaries was combined with data collection by a participant observer (both fieldnotes and tape recordings) in a doctoral dissertation by Brown (1983). The most accessible version of this study is found in her (1985a) paper on the input addressed to older and younger learners of Spanish as a foreign language. Her subjects were eighteen younger learners (ranging in age from 19 to 23, with a median age of 20.9) and eighteen older learners (between 55 and 75 years of age, with a median age of 63.6 years). Both groups included male and female students. All thirty-six learners were enrolled in an intensive Spanish program at the Mission Training Center at Brigham Young University, where they were in class six hours per day and had an additional two hours per day of homework. Brown's reason for studying these groups was that the older learners in the Mission

Training Center (MTC) had typically experienced more difficulty than the younger learners in trying to master Spanish.

The learners in Brown's study were given these instructions on keeping their diaries (1985a:283-284):

> "The journal has two purposes. The first is to help you with your language learning. As you write about what you think and feel as a language learner, you will understand yourself and your experience better.
>
> The second purpose is to increase the overall knowledge about language learning, so that learning can be increased. You will be asked to leave your language learning journal when you leave the MTC. However, your journal will not be read by teachers at the MTC. It will be read by researchers interested in language learning.
>
> Your identity and the identity of others you may write about will be unknown (unless you wish it otherwise) to anyone except the researchers.
>
> You will be given 15 minutes a day to write. Please write as if this were your personal journal about your language learning experience."

Brown analyzed the resulting diaries for "any reference to input desired, to amount of input given; to type, complexity or meaningfulness of input" (1985a:278). She found a different focus on the amount of input the learners were receiving: The younger learners wrote about input four times as often as the older learners, and this difference was statistically significant. (As far as I know, Brown was the first person to use inferential statistics in analyzing learners' diaries.) While there were only minimal quantitative differences in the learners' comments about the types of input they received, Brown notes a qualitative difference in the learners' diary entries: Over 28.2% of writing by the older learners dealt with desired changes in the input. In contrast, the younger learners' writing suggested changes only 2.7% of the time. The older learners' requests for instructional alteration may be indicative of their discomfort, but Brown does not speculate on this point.

The strengths of Brown's work are numerous. Hers is the first diary study to use comparison groups (in a criterion group design) and the combination of statistical and qualitative analyses of qualitative data. In addition, Brown's (1985b) use of both participant observation fieldnotes and learners' diaries permits

triangulation (Denzin, 1978; Fry, 1988; van Lier, 1988). The concept of triangulation has been borrowed from surveying, as Hammersley and Atkinson (1983:198) explain:

> "For someone who wanted to locate their position on a map, a single landmark can only provide the information that they are situated somewhere along a line in a particular direction from that landmark. With two landmarks, however, their exact position can be pinpointed by taking bearings on both landmarks; they are at the point where the two lines cross. In social research, if one relies on a single piece of data, there is a danger that undetected error in the data-production process may render the analysis incorrect: If, on the other hand, diverse kinds of data lead to the same conclusion, one can be a little more confident in that conclusion."

Brown (1985b) describes how she used learners' diaries and audio recordings of class sessions and her ethnographic fieldnotes to corroborate the inferences she drew.

In a brief section of a paper with a broader focus, Grandcolas and Soulé-Susbielles (1986) report on the use of diaries kept by French teachers-in-training who were studying language as part of their professional preparation. (This article does not mention the number of diarists involved.) The reported findings include (1) the importance of the teacher's personality and attitude; (2) the role played by the diarist's student peer group; and (3) the necessity of the personal commitment of the language learner. A fourth factor, which these authors call "shifted enunciation," refers to the fact that even when students interact with other students (e.g., when one student asks a question of his neighbor), the teacher is still the intended audience of the communication. Grandcolas and Soulé-Susbielles interpret this factor as underlining the importance of teacher-learner relationships, as opposed to learner-learner relationships. Unfortunately, in this publication the authors have space to provide only a brief report of the diary project. Interested readers are referred to Grandcolas (1986).

In 1989, Matsumoto reported on the language learning experience of a nineteen-year-old Japanese girl (called "M"), who attended an intensive English program in the U.S. for eight weeks in the summer. Matsumoto conducted a frequency count of factors mentioned in the thirty-six entries in M's diary, and reported the results as percentage data. This process identified nine learning activities, nine clusters of emotional factors, and sixteen "non-emotional factors"

which appeared in M's journal and are interpreted by Matsumoto as important to the language learning process. The journal entries, some of which are included in the article, were supplemented by questionnaire and interview data. Matsumoto also compared M's issues with those discussed in other diaries.

Ellis (1989a) also used learners' diaries along with other data sets in his analysis of classroom learning styles and their effect on second language acquisition. In what is probably the strongest data triangulation effort to date, Ellis utilized information from questionnaires, cognitive style testing, a language aptitude test, attendance and participation records, a word order acquisition score, speech rate and three proficiency tests, in addition to the journals of two adult learners, Simon and Monique, as they took a beginning German course at a college in London. The course was described as being "almost entirely form-focused" with instruction "fairly evenly divided between practise and consciousness-raising activities" (Ellis, 1989a:251). The goal of the course was "to develop a high level of linguistic accuracy in the use of L2 German" (ibid.:252). Regarding his procedures, Ellis wrote, "The learners kept journals of their reactions to the course, their teachers, their fellow students, and any other factors which they considered were having an effect on their language learning" (ibid:252-253).

Ellis identified four key variables in the second language acquisition research literature on good language learners. In his review (see Ellis, 1990) the following factors were consistently related to effective language learning:

(1) a concern for language form;
(2) a concern for communication;
(3) an active task approach; and
(4) awareness of the learning process.

Ellis reports that Monique's and Simon's diaries provide "ample evidence" of the extent to which the two learners manifested these traits. Simon's diary documented all four, but Monique's revealed a lack of concern for communication in coping with the German course. Her extreme focus on form to the exclusion of an emphasis on communication is curious, since she was a native speaker of Creole but spoke both French and English fluently and accurately, and had lived in a multi-lingual society and used all three languages for communication (1989a:251). Ellis interprets this pattern as an adaptive response to the formal instructional context, and possibly a subordination of Monique's natural language learning patterns, as revealed in her history and other data sets. Ellis notes that "Monique's cognitive orientation was

almost entirely studial. Her journal shows that she is obsessively concerned with linguistic accuracy" (ibid.:254).

Judging from her performance on the formal outcome measures, Monique did relatively well in her coursework and met the accuracy goals of the German class. However, she performed poorly on an oral narrative task designed to assess fluency. On the accuracy measures, her scores are equal to or higher than Simon's. It is only through Monique's diary entries that we see how uncomfortable she was with the formal emphasis of the course. Ellis (1989a:257) reprints this and other direct quotes from Monique's diary:

> "I was quite frightened when asked questions again. I don't know why; the teacher does not frighten me, but my mind is blocked when I'm asked questions. I fear lest I give the wrong answer and will then discourage the teacher as well as be the laughing stock of the class maybe. Anyway, I felt really stupid and helpless in that class."

Ellis concludes that Monique probably would have benefited from a comprehension-based approach to learning German in the initial stages. He writes (ibid.:258-259),

> "Monique does not appear unduly disadvantaged, as she performs well in the grammar proficiency test -- i.e., she succeeds in developing the grammatical accuracy needed to succeed in the course. However, she pays a price. The course proves a painful experience and she is unable to perform effectively in a communicative task."

In spite of her discomfort, Monique attended 96% of the class sessions, compared to 86% for Simon, but Simon took more in-class practice opportunities than Monique. This brief but fascinating paper left me wondering how many low-level language courses are designed for students like Simon, but peopled with learners like Monique.[7]

These, then, are the published diary studies I have located in which the data were analyzed by researchers other than the diarists themselves. Table 2 lists these studies and summarizes information on the learners and the languages involved:

Table 2

Diary Studies in Which Researchers Analysed Other Learners' Diary Texts

AUTHOR(S)/DATE(S)	LEARNER(S)	TARGET LANGUAGE(S)
Matsumoto (1989)	1 19-year-old Japanese learner	English
Bailey (1983)	11 learners, mostly teachers and linguists (though Fields' & Moore's reports were also used)	Several target languages
Parkinson & Howell-Richardson (1990); Howell-Richardson & Parkinson (1988)	51 foreign learners of English in Scotland; an unknown number of Scottish people learning foreign languages	English; French, Spanish and other unspecified target languages
Brown (1983, 1985a and 1985b)	36 older and younger learners	Spanish
Grandcolas and Soulé-Susbielles (1986)	Unknown number of teachers-in-training	French*
Ellis (1989a)	2 adult learners	German

*(I infer that French is the language under study, since the learners were French teachers-in-training.)

The Doubting Game

Given this review of the language learning diary studies, we can now return to the doubting game and question the value (indeed, the premises) of this approach to SLA research. The doubting game is based on skepticism: "Doubting an assertion is the best way to find an error in it" (Elbow, 1973:148). Playing the doubting game is not <u>just</u> criticizing, but it does involve taking a critical stance.

Although I do not wish to equate the doubting game and experimental science, it is especially easy to play the doubting game with regard to the diary studies if one has been trained in the experimental paradigm. From this perspective, concerns

about the diary studies can be divided into three main categories: (1) problems regarding the subjects, (2) problems in data collection, and (3) problems in data analysis.

Many of these issues are related to the concept of generalizability as a desired goal of experimental research. Generalizability (or external validity, as it is also called) is the extent to which the findings of a study can be applied beyond the context of the original investigation. In the experimental paradigm, hypotheses are tested on a sample of subjects, carefully selected to represent the broader population of interest. An inherent assumption in this form of research is that the findings of such studies can be generalized to that population.

Problems Regarding Subjects

The usual concerns regarding subjects in the diary studies hinge around the small number of learners involved. As with other case studies, there is often only one subject (see, e.g., Bailey, 1980; Danielson, 1981; Fields, 1978; Jones, 1980; Matsumoto, 1989; Moore, 1977; Rivers, 1979 and 1983; Rubin and Henze, 1981; Schumann, 1980). Obviously a learner can only introspect about his or her own learning processes (by definition), but it is also possible to compare the findings from different learners' diaries. Ellis (1989a) and Schumann and Schumann (1977) both reported on two learners, while Bailey (1983) reviewed eleven learners' journals (or the published reports based upon them). Journals from thirty-six learners of Spanish were studied by Brown (1983, 1985a, 1985b). The largest number of subjects reported to date was the cohort of 51 learners of English in the study by Parkinson and Howell-Richardson (1990). But even fifty-one is a very small number by experimental standards. The concern about large subject pools in experimental research is that (1) large samples are more representative of the population, and (2) most statistical procedures work more reliably with large numbers of subjects.

Another concern about several of the diary studies published thus far is the fact that many of the diarists were themselves linguists, experienced teachers, or language teachers-in-training. This is the case in all the published diary studies except for Brown (1983, 1985a, 1985b), Ellis (1989a), Fields (1978), Matsumoto (1989), Moore (1977), and Parkinson and Howell-Richardson (1990). As Parkinson and Howell-Richardson (1990:128) note, in these instances

> "the diarists are linguists...going 'back to school' after (and usually during) work as teachers/researchers/teacher educators, and their perceptions are inevitably affected by this: They sometimes find their theoretical sophistication surprisingly unhelpful, but it cannot fail to colour their report."

The concern is that these researchers/diarists may not be typical of other language learners -- either because they are potentially better language learners or because they may have more metacognitive or metalinguistic awareness. (According to Mann, 1982:89, metacognitive awareness means the subjects can "observe the contents of their minds and infer from this observation the processes in operation.") Teachers as language learners may also experience more ego-involvement in the language learning process than would other language learners. For these reasons, the concern is that what we learn about language teachers or linguists as language learners may not be generalizable to other students.

Problems in Data Collection

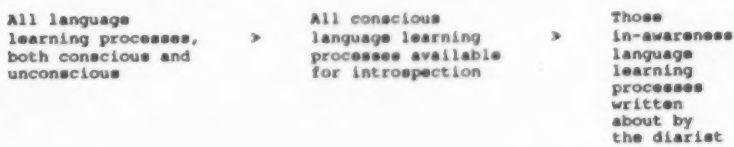
Another problem area for the diary studies involves the data collection process. By definition, the diary data are produced by the learner himself, recording, reacting to, and reflecting upon his experiences. Thus the diary studies are subject to all the difficulties associated with other forms of self-report (Oller and Perkins, 1978) and introspective and retrospective data collection (see Faerch and Kasper, 1987; Grotjahn, 1987; Seliger, 1983). The diaries involve subjective data, based entirely on the learners' perceptions of their experiences. They are by no means objective reports, and objectivity is one of the desired hallmarks of experimental research.

Furthermore, in cases where the journal entries are made after the event (immediate to delayed retrospection), there are serious concerns about how much time elapsed between the event and the recording (Seliger, 1983). As Fry points out (1988:160) "With retrospective data, all the problems of cognition in introspective data are magnified by the lapse of time between the event and the reporting of it."

Another serious concern about data derived through introspection is the question of what parts of our mental processing are actually open to examination. Seliger's (1983) methodological review examines, among other things, "psychoanalytic" studies utilizing "mentalistic data." He categorizes the diary studies and work involving think-aloud protocols within this type of research.

Seliger points out that data from the "psychoanalytic" studies can be evaluated in two ways: (1) in terms of what they tell us about the affective domain, and (2) "for what they can tell us about the processes of language learning itself" (ibid.:187). The problem with the latter, of course, is that we do not know how many or which of these language learning processes operate within learners' conscious awareness and are therefore available as objects of introspection. In the case of the diary studies, those language learning processes which learners actually choose to write about are potentially a smaller group than are all the conscious processes which learners might write about, and this subset of conscious processes is presumably smaller than the entire range of language learning processes, both conscious and unconscious, which influence second language acquisition. This relationship is depicted in Figure 2:

Figure 2: Subsets of Language Learning Processes


All language learning processes, both conscious and unconscious
All conscious language learning processes available for introspection
Those in-awareness language learning processes written about by the diarist

Even if we reject outright Seliger's (1983:187) claim that "obviously, it is at the unconscious level that language learning takes place," we must acknowledge that the diaries, as data collection devices, can only access some (as yet unspecified) subset of all language learning processes.

Likewise the quality of the journal entries varies from "thick description" (Geertz, 1973:6) to sketchy reports. If the diarist records only externally verifiable facts, we are left with a flat account, which could have been derived more reliably from transcribed tape recordings of the event. On the other hand, if the diarist records only reactions, without detailing the events leading to these reactions, the reader is left with assertions that lack credibility. Fry (1988:161) states that "consistency, in terms of time (i.e., that the data is recorded at a fixed time after the event, preferably as soon as possible) and in terms of depth (i.e., the level of detail

recorded) cannot be ensured." Seliger also notes that most language learners are not trained linguists and do not, therefore, have available the means for describing linguistic processes or for interpreting them reliably. And, if the diarist revises the journal for public consumption, as described in Bailey and Ochsner (1983), we have no way of knowing how much information has been deleted or changed.

Finally, as Fry points out, the diary-keeping process itself is extremely time-consuming "and initial enthusiasm may give way to fatigue" (1988:161). Rubin and Henze (1981:17) claim that "the amount of time spent keeping a diary can be reduced if students are directed to focus on specific aspects of cognitive learning." As a time-saving device, these authors suggest that learners make notes and mark them with an asterisk in their regular class notebooks (ibid.). Brown (1985b:132) claims that "both participant observation and diary studies require considerable time writing up the data and analyzing them, but participant observation takes more time in gathering the data." Nevertheless, as with any longitudinal case study, keeping a journal over a long period of time demands a commitment on the part of the diarist.

The issue of time, in turn, raises questions about the data collected in the non-introspective diary studies (i.e., those in which a researcher analyzes diary entries made by other language learners). What is the commitment, and what sort of quality can we expect from diarists who are required, in some sense, to keep intensive journals? Both Asher (1983) and Brown (1985b) discuss the difficulties of gathering data from learners, and the variable quality of data from different informants is a much-discussed problem in ethnography (see, e.g., Georges and Jones, 1980; Hammersley and Atkinson, 1983; Johnson, 1975; and Malinowski, 1989).

Problems in Data Analysis

Problems in data analysis are related both to concerns about the subjects and about the sort of data generated by the diaries. Given the small number of subjects, and the fact that many have been potentially atypical second language learners, the generalizability of the findings of such studies is limited. Furthermore, with the lack of control over variables and the pre-experimental nature of the design, causal statements are not possible.

The data generated by the diarists are also subject to all the problems associated with other types of qualitative data analysis. These include issues of data reduction

(whether through summarizing, coding, developing typologies, etc.), the definition of categories, the open-ended nature of the data, reliability in coding and interpretation, etc. We must also be aware that the published diary studies are either extremely long (e.g., Schmidt and Frota, 1986, is ninety pages) or may involve abstractions and inferences with no supporting data (e.g., Schumann and Schumann, 1977). While quantitative data can be easily summarized, through the conventions of descriptive statistics, and presented in tabular form for economy of reporting, it is difficult to convey the compelling contents of learners' diaries without quoting excerpts from the journals themselves.

Part of the difficulty here is that second language acquisition research involving interpretive analyses of qualitative data lacks what Kaplan (1964:3-11) has called "reconstructed logic," instead, we have mainly "logic in use" (ibid.). In contrast, the familiar statistical operations associated with the experimental paradigm (i.e., the quantitative analysis of typically quantitative data) provide researchers working in that tradition with clear-cut procedures for making decisions about statistical significance. Experimental science and statistical procedures also involve a good deal of "logic in use," but this fact is often obscured by the clearly delineated "reconstructed logic" of the experimental paradigm.[8]

In analyzing the journal entries, the researcher (and subsequently the reader) must ask, "What constitutes a pattern? What makes an event 'salient' to the learner? How are key terms defined?" Such methodological issues have been widely discussed in qualitative approaches to data analysis in sociology (e.g., Brown, 1977; Johnson, 1975; Krippendorf, 1980; Morgan, 1983; Reason, 1988;), education (Marshall and Rossman, 1989), evaluation (Cook and Reichardt, 1979; Guba, 1978; Patton, 1980), and anthropology (e.g., Dobbert, 1982; Georges and Jones, 1980; Spradley, 1979, 1980). Likewise the life history approach (see Bertaux, 1981, and Watson, 1976) is widely used in social science research, as is the case study method (Yin, 1984). However, the interpretive procedures for analyzing qualitative data have not yet been well codified in the methodology literature on second language acquisition research, with the possible exception of discourse analysis. (An exception is van Lier, 1988, which deals with both coding and transcription.)

Furthermore, the diary studies have apparently not born fruit in terms of early claims about their potential usefulness as hypothesis-generating tools, although Fry considers the hypothesis-generating role to be the "most tenable of the claims" made about diary studies (1988:164). It is not clear whether this gap is a result of the diary studies themselves failing to suggest testable hypotheses, or simply the

result of experimental researchers choosing not to utilize the constructs or hypotheses suggested by such studies.

In short, from the point of view of experimental research, there are numerous problems associated with the diary studies, operating at all levels (of subjects, data collection, and data analysis). Most of these problems hinge around the notion of generalizability, the precept that the results of experiments or ex post facto studies on samples should be generalized to the wider population. The diary studies fall short of most of the requirements of external validity, and are therefore potentially uninteresting to researchers trained in the experimental paradigm.

The Believing Game

But achieving generalizability is neither the purpose nor the point of the diary studies. As van Lier points out (1988:2-3):

> "There has been almost unanimous pressure to choose topics for research that can be readily generalized to larger populations.... We are all agreed that greater understanding of language learners is also a legitimate activity."

The point of the diary studies is to understand language learning phenomena and related variables from the learner's point of view. If we set aside the notion that generalizability is the sole legitimate focus of SLA research, it then becomes very easy to play the believing game with regard to the diary studies.

Why should we play the believing game? Elbow, the author of the metaphor, claims that

> "By believing an assertion we can get farther and farther into it, see more and more things in terms of it or 'through' it, use it as a hypothesis to climb higher and higher to a point from which more can be seen and understood" (1973:163).

This is exactly what Larsen-Freeman did (in 1981), when she played the believing game with four different models of second language acquisition, trying to see what each one could reveal to her about the role of cognition in SLA, if she accepted the model's premises and arguments. We will play the believing game, with regard to

the language learning diary studies, in terms of what they have to offer teachers, learners, and second language researchers.

Benefits for Language Teachers

The experience of studying a language again, when one is already a language teacher, is always revealing. But the powerful and sometimes surprising insights one gains by struggling with a new linguistic system and all the emotional baggage that goes with it can best be captured and later reflected upon by keeping a diary. Rubin and Henze (1981) note that in pre-service training, the typical language requirement can be enhanced by having the teachers-in-training keep a journal of their experiences and analyze the trends. For instance, Danielson, an experienced teacher and teacher trainer, reports re-experiencing things which she knew but had lost touch with. I purposefully use this tactile expression to match the kinesthetic and visual images used in Danielson's entry:

> "Once again, I was actually experiencing and reflecting on many things which I intellectually understood but no longer felt or saw from a learner's point of view and many things which I had long ago tucked neatly away and forgotten" (1981:16; underscoring added).

Bailey (1983:76) discusses how the imagistic language of the diary entries can reveal the learner's attitude. Teachers reading learners' journals (and researchers who analyze such journals as data) should be aware of the metaphors students select to express themselves. (For a similar perspective in the teacher supervision literature, see Gebhard, 1984:509-512.)

Even without keeping journals as learners, teachers can benefit greatly from reading the available diary studies. The account of Simon and Monique's different reactions to their German program cannot help but remind us that our classes are full of dissimilar students, whose needs are only more or less met by our instructional decisions (Ellis, 1989a). Anyone working with older learners should read Brown's (1985a) and Danielson's (1981) reports. Teachers struggling with issues of testing and grading should read Bailey's (1980) account of students' reactions to an "unfair" test. And Schmidt's frustrations with his teacher denying the truth value of his utterances for the sake of the drill will sound uncomfortably familiar to instructors who have tried to balance the goals of the lesson with the learners' emerging personal agendas. These vignettes from the learners are

powerful and sometimes disquieting reminders of the students' central role in classroom second language learning, and of how seldom we as teachers really know their points of view.

Benefits for Language Learners

How can the diary studies benefit language learners, other than by making teachers more aware of and sensitive to the learners' perceptions, strategies, and feelings? It is my belief that it is useful both for learners to keep journals and to read diary studies by and about other language learners. First we will consider the benefits of keeping journals.

The frustrations of learning a second language are well documented in SLA research. Some authors feel that keeping a diary provides a safety valve. This is the cathartic function of diary studies: Learners may write out their frustrations instead of skipping or dropping class, or harboring grudges against the teacher or other learners (Bailey, 1983; Bailey and Ochsner, 1983:193).

Some learners have the perception that keeping a language learning diary can promote awareness of second language learning processes and pitfalls. The following comments were taken from learners' diaries cited in Grandcolas and Soulé-Susbielles (1986:301):

> "This observation work has made me aware of the part I was able to play in my learning of the language. It was possible for me to take part actively in this course, even with this traditional method, as I really wanted to learn something else differently. It seems to me that, if every learner was made aware of his/her learning, the development of the course would be changed."
>
> "Our part of learner/observer has made our utterances less spontaneous, more consciously organized for the checking of such or such a hypothesis.... We were much more sensitive to what was important and what was not.... Let us not go to a language lesson as passive consumers!"

These sentiments are echoed by Rubin and Henze, in their co-authored analysis of Henze's journal (1981:24):

> "There are definite benefits from making such observations. Henze (the learner) said that the research helped her to focus her learning, and that by the end of the course she had concretized some vague notions about her own learning by providing specific examples. After the study, Henze could more clearly see how she uses her knowledge of other foreign language structures in the comparison and modification of hypotheses in learning Arabic. In addition, the diary helped Henze to evaluate her own learning strategies enabling her in some cases to manipulate situations so that she received the most benefit."

In this case, the authors claim, both making the diary entries and later analyzing them were helpful to the teacher-turned-learner.

Brown also found that, in the diary entries, learners were able to recognize their own progress and suggest ways to improve the instructional program. She states that many learners "gave evidence in their journals of being aware of their progress. It may be that the awareness would have come without the journals, but writing it down made it very evident" (1985b:131).

As Fry has noted (1988:161), "the act of recording aspects of learning behaviour will raise consciousness of that behaviour and may change it." This is, of course, a restatement of Labov's (1972) classic "observer's paradox." But this fact, which is a potential drawback for researchers, can be a tremendous asset for language learners. Asher (1983) has documented ways that she used published diary studies in helping to make adolescent learners of French more aware of their own learning strategies. There are probably several ways that diaries could be used by creative teachers in learner training programs.

Benefits for Language Learning Research

As noted above, diary studies of language learners in formal instructional settings are part of the emerging tradition of language classroom research, as described by Allwright (1983), Allwright and Bailey (1991), Bailey (1985), Brumfit and Mitchell (1990), Chaudron (1988), Gaies (1983), Long (1980, 1983), Mitchell (1985), and van Lier (1984, 1988, 1989). With data generated by the learners themselves, the diary studies provide us with views from "inside the 'black box'," to use Long's (1980, 1983) metaphor for the unexplored processes of classroom language learning. Like other introspective methods, diary studies give us

information about the language learners and particularly about their perspectives on affective and instructional factors which influence second language learning.

The use of diaries as a source of data is a well established procedure in naturalistic inquiry (e.g., see Hammersley and Atkinson, 1983; Malinowski, 1989; Plummer, 1983:17-21; and Yin, 1984). While the diary studies are not experimental, they are empirical, in the sense that they start with the collection of data -- data, in Bateson's (1972) sense as "records of events." The learners' journals, however, provide both records of events and, through introspection and retrospection, the learners' responses to those events. In an age when both pedagogy and curriculum development have recognized the learner's central role, it is appropriate that researchers should also bring the learner into the picture. In fact, Allwright and Bailey (1991) have suggested ways for classroom researchers to include learners collaboratively in the research process. The lingering question is, what do diary studies have to offer research?

I would argue that the diary studies are absolutely essential to advancing our understanding of classroom language learning. At the present time we are working with an unrefined tool to craft an only dimly understood representation of language learning. Properly done, the diary studies can provide us with important missing pieces in this incredibly complex mosaic -- pieces which may not be fully accessible by any other means.

Diary studies allow us to see factors identified by the learners which we, as researchers and teachers, may not consider to be variables worth studying. The lack of researcher control over variables, which is seen as a problem in experimental science, is viewed as a strength of the naturalistic inquiry tradition. Experimental research on classroom language learning has often been criticized for its laboratory-like cleanliness and disregard for context (e.g., van Lier, 1988). One strength of the diary studies to date is that they reflect the "real-world" conditions under which the data were collected: F. Schumann writes about the discomforts of her living quarters in Tunisia; Schmidt relates his extreme annoyance with the Brazilians who talk about him in Portuguese, not knowing that he understands; Jones expresses her humiliation in the Indonesian class, where the program director belittles her in front of the other students. These are all factors which are not controlled in experimental research -- nor are they usually even addressed. But for these learners, and presumably for many others, they were powerful negative experiences which were perceived as having an influence in the language learning process.

Another positive factor is that the primary data collection process, keeping a diary, is "low-tech," portable, and trainable. It can be used by any language learner with the will to introspect and retrospect, and the ability to keep writing. It does not require extensive preparation in test development, questionnaire design, or statistical procedures. Nor is it expensive in terms of equipment and materials.

This does not mean, however, that anyone can and should conduct diary studies. Rather, the basic prerequisites for data collection include at least a willingness to view oneself critically, and the ability to question one's motives and write comfortably and consistently, regularly and at great length, without premature editing. This is a very difficult task indeed. (See Yin, 1984:55-60, for a more detailed discussion of desired skills for case study researchers.)

Furthermore, the diaries can provide valuable sources of data triangulation (Denzin, 1978; Fry, 1988; van Lier, 1988) when used with other sources of data. This strength has been aptly demonstrated by Brown (1983, 1985a, 1985b), Ellis (1989a), Matsumoto (1989), and Schmidt and Frota (1986). Given what we are beginning to learn about individual differences in language learning (e.g., Ellis, 1989b) SLA research which fails to take personal variables into account must be interpreted very cautiously. Without the diaries written by Simon and Monique, Ellis (1989a) would have had an impressive battery of test results on the two learners, but no way of accessing their very different responses to their instructional program.

An appreciation for individual differences (which Schumann and Schumann, 1977, called "personal variables") leads directly to the importance of studying single learners in depth. The value of the detailed case as an exemplar has long been recognized in the life history approach (e.g., Bertaux, 1981; Watson, 1976) and in case study research (see Yin, 1984, for a thorough methodological discussion). The language learning case studies collected by Hatch (1978) yielded extremely important ideas in the early days of second language acquisition research.

It is certainly useful, in the search for generalizable findings, to obtain measures of central tendency and measures of dispersion of scores about the mean in a group of scores, but these data and the subsequent inferential statistics which are conducted with them do not supply all we need to know about language learning -- particularly in classrooms, where there is at least some obligation to help the learners get on with learning. If we had only means and standard deviations and tests of statistically significant differences, we could not tell how profoundly and

distressingly different Monique was from Simon (in Ellis, 1989a). Nor can we begin to understand the factors which drive people from the language classroom unless we listen to the learners: the drop-outs, the discouraged and the overwhelmed who often just disappear from experimental studies, or suffer through the course to the end without our discovering why they did poorly (or even well) on (some of) the dependent variables. Here again we find a parallel between the doubting game and experimental research: "The doubting game deals with classes of things...whereas the believing game deals with particular, unique things" (Elbow, 1973:165). SLA research, in its zeal to generalize, must not lose sight of the individual learner.

It has been argued by Fry (1988:163-164) that the diary studies "have revealed nothing that directly contributes to our understanding of SLA processes." Fry's review, however, does not include some of the more substantial diary studies (e.g., Bailey, 1983; Brown, 1983, 1985a, 1985b; Ellis, 1989a; and especially Schmidt and Frota, 1986) which are now available. (I think Fry's claims are essentially right for the two diary studies he does cite: Bailey, 1980, and Schumann and Schumann, 1977).

As we improve the diary study tool, the resulting SLA findings will also be more helpful. For instance, in Schmidt and Frota (1986), the combination of Schmidt's linguistic insights and Frota's native-speaker awareness leads to very convincing comments about second language acquisition. Even though Krashen (1983) introduced the "notice-the-gap" principle, it is now Schmidt and Frota's revision of that principle, supported by illustrations from "R's" journal, which is most often cited in the SLA literature. Likewise, Schmidt's experiences, as documented in his journal, shed new light on the autoinput hypothesis, which was suggested in an earlier form by Gregg (1984) and Sharwood-Smith (1981), as Schmidt and Frota point out. Such diary studies have the capacity to add depth, detail, and realism to existing hypotheses and theoretical constructs.

The diary studies are also ready sources of illustrations for researchers to use with non-researchers. As Peter Shaw has pointed out (personal communication, 1991), having clear prose examples of language learning as perceived by the learners themselves can be extremely useful for communicating with audiences who are not trained in the interpretation of statistical reporting. Such audiences include many teachers, most language learners, some funding agency representatives, parents, administrators and the media.

Finally, learners' journals and the resulting diary studies can offer researchers a wealth of new ideas and questions about second language learning. The following come immediately to mind:

(1) Following from Parkinson and Howell-Richardson's (1990) discussion of learners' perceptions of time spent on language learning, van Lier (personal communication) has suggested we consider the role of quality time (as opposed to quantity of time) as a variable in SLA.

(2) Given the cooperation/competition factor which has emerged in some diary studies (e.g., Schumann, 1980; Bailey, 1983), how can couples (Brown, 1985a; Schumann and Schumann, 1977) and teachers or groups of classroom learners manipulate this factor to enhance language learning?

(3) For learners in second language situations, what are the respective roles of in-class instruction and out-of-class interaction in promoting and/or inhibiting second language learning? (See Jones, 1977, and Schmidt and Frota, 1986.)

(4) What do learners gain in situations where the variable linguistic input available is sometimes more and sometimes less demanding and challenging? (See Danielson, 1981, and Schmidt and Frota, 1986.)

(5) How do the language learning experiences of non-linguists (e.g., Fields, 1978, and Moore, 1977) differ from those of linguists, if at all?

(6) What language classroom factors lead to debilitating anxiety (Bailey, 1983; Parkinson and Howell-Richardson, 1990) and how can such anxiety be managed?

(7) What differences emerge in the perceptions of learners in the same class, program, or situation (e.g., Monique and Simon in Ellis, 1989a; Schumann and Schumann, 1977)?

These questions (and many others) can be addressed by diary studies, provided they are done well.

Suggestions for Further Reading

Other sources of information may be useful to people interested in using this approach to research. These include the methodological papers which have been written about the diary studies. Both Matsumoto (1987) and Fry (1988) have published critical reviews of diary studies in the journal of the Japan Association of Language Teachers. Matsumoto's work is generally laudatory: She plays the believing game. Fry's work is more critical: He plays the doubting game through most of his article, but then comments on the usefulness of diary studies in action research (Kemmis and Henry, 1989; Kemmis and MacTaggart, 1982; Nunan, 1989, 1990). Brown (1985b) has compared the diary studies and participant observation as two approaches to language classroom research, and Howell-Richardson and Parkinson (1988) have discussed the "possibilities and pitfalls" of learner diaries. Bailey and Ochsner (1983) make a case for improving the quality of diary studies by (1) establishing believability, (2) identifying (and identifying with) the audience, (3) minimizing author distance, (4) providing information about the second language learning context, and (5) explaining in some detail how the diary data were collected and analyzed.

Other related papers are concerned not with diary studies per se but with a broader focus on introspective and retrospective data collection in general. These include Cohen and Hosenfeld's (1981) paper on uses of mentalistic data in SLA research, which is generally positive. The volume by Faerch and Kasper (1987) -- particularly the methodological paper by Grotjahn -- provides a valuable treatment of introspective methods. An important criticism of this type of work is by Seliger (1983), who argues that learners' "verbal reports can be taken as a starting point for research, not as an empirical conclusion" (1983:185). Chaudron's (1983) article on metalinguistic judgments is helpful, and Mann (1982) offers numerous useful suggestions for improving the quality of introspective data -- particularly the think-aloud form of verbal protocols. Working in a different tradition, Churchland (1990) discusses introspection and its connection to conscious knowledge.

Numerous references to classroom research have already been cited. As part of a broader research tradition, language learning diary studies are part of naturalistic inquiry (see Guba 1978; Lincoln and Guba, 1985), and are most closely related to ethnography. For articles on ethnography in language-related research, see Watson-Gegeo (1988) and van Lier (1990). Books on this topic include Cazden (1988), Saville-Troike (1982), and van Lier (1988). The collections edited by Green and Wallat (1981) and Trueba, Guthrie and Au (1981) are useful anthologies.

Information dealing specifically with quantitative and qualitative approaches to second language classroom research can be found in Chaudron (1986). In the broader arena of general educational research, Cook and Reichardt (1979) have edited a book on qualitative and quantitative research methods. Yin's (1984) book on case study research would be very useful to anyone considering doing a diary study.

For people wishing to read more about competing research paradigms, an excellent starting point is Ochsner's (1979) paper, which argues for a bilingual perspective in SLA research. He discusses the history of both the nomothetic (experimental) paradigm and the hermeneutic (interpretive) tradition, in which the diary studies may be categorized. Ochsner's work influenced both Long's (1980, 1983) ideas on approaches to language classroom research, and Schumann's (1983) discussion of art and science in SLA research. Smith and Heshusius (1986) outline the history of relations between positivistic (experimental) science and naturalistic inquiry, highlighting the major philosophical differences between the two approaches.

In this paper, I have tried to define the diary studies and to locate them within classroom research in the naturalistic inquiry tradition. Next I summarized the findings of the diary studies published to date. Then I tried to raise concerns by playing "the doubting game," and to raise interest by playing the "believing game." My intent was to play both games equally well, but I believe that it is of potentially more importance to play the doubting game from the perspective of one who accepts the genre (and can therefore bring an appropriately critical eye to bear upon its offerings to date) than to play just the believing game. As van Lier has pointed out (1989:42),

> "The blacksmith cannot criticize the carpenter for not heating the piece of wood over a fire. However, the carpenter must demonstrate a principled control over the materials used."

I hope that in this paper I have demonstrated that the diary studies are at least an emergent craft. We have not perfected the use of learners' diaries as tools. Indeed, only a few of the published diary studies (e.g., Schmidt and Frota, 1986) have been exemplary. Much more work needs to be done.

Endnotes

1. I would like to thank Peter Shaw, Leo van Lier, Cherry Campbell, and Ruth Larimer for their perceptive comments on earlier versions of this paper. My thanks are also due to Amy Saviers, Naomi Kubota Fujishima and Vicki Voll for their hard work at the word processor: Coping with the numerous generations of this manuscript has shown them the dark side of process writing.

2. I apologize if I have overlooked any published diary studies and would appreciate being informed of the citations.

3. Asher also conducted directed diary research when she supervised a group of eight adolescents receiving French instruction in the U.S. and then subsequently using their French in a homestay program in Switzerland. Asher gave her learners several different tasks to do, related to discovering their own learning patterns. These tasks included reading and analyzing diary entries from Bailey (1980) and Jones (1980). More information can be found in Asher's unpublished masters thesis (1983).

4. It would probably be worthwhile for someone considering doing a diary study to attempt an analysis of River's (1979, 1983) data, as a way of getting familiar with the process.

5. Although we do not often make the connection in present-day English, competent and compete originally derived from the same root words. (See the Oxford English Dictionary, 1971:718-719.)

6. I apologize to the authors for not having discussed Howell-Richardson and Parkinson (1988) in more depth. It was not available to me as I wrote this article.

7. Following Asher (1983), I believe learners' diaries can be useful tools in learner training programs. The study by Ellis (1989) is accessible and clear, and the contrast of the two learners' styles would provide useful discussion material for other second language learners to read.

8. It is not my intent here to criticize experimental science or to point out the numerous short-comings of working with quantitative data: When I wish to determine causality or correlations, I work with those research designs and the appropriate statistical tools.

REFERENCES

ABRAHAM, R.G. and R.J. Vann. 1987. Strategies of Two Language Learners: A Case Study. In A. Wenden and J. Rubin (eds.), Learner Strategies in Language Learning. Englewood Cliffs, NJ: Prentice-Hall International, 85-102.

ALLWRIGHT, D. and K.M. Bailey. 1991. Focus on the Language Classroom: An Introduction to Classroom Research for Language Teachers. Cambridge: Cambridge University Press.

ALLWRIGHT, R.L. 1983. Classroom-centered Research on Language Teaching and Learning: A Brief Historical Overview. TESOL Quarterly 17(2): 191-204.

ASHER, A.L. 1983. Language Acquisition Diaries: Developing an Awareness of Personal Learning Strategies. Unpublished Master's Thesis. School for International Training, Brattleboro, Vermont.

BAILEY, K.M. 1980. An Introspective Analysis of An Individual's Language Learning Experience. In R. Scarcella and S. Krashen (eds.), Research in Second Language Acquisition: Selected Papers of the Los Angeles Second Language Research Forum. Rowley, Mass: Newbury House, 58-65.

BAILEY, K.M. 1983. Competitiveness and Anxiety in Adult Second Language Learning: Looking At and Through the Diary Studies. In H.W. Seliger and M.H. Long (eds.), Classroom Oriented Research in Second Language Acquisition. Rowley, Mass: Newsbury House, 67-103.

BAILEY, K.M. 1985. Classroom-centred Research on Language Teaching and Learning. In M. Celce-Murcia (ed.), Beyond Basics: Issues and Research in TESOL. Rowley, Mass: Newbury House, 96-121.

BAILEY, K.M. 1990. The Use of Diary Studies in Teacher Education Programs. In J.C. Richards and D. Nunan (eds.), Second Language Teacher Education. Cambridge: Cambridge University Press, 215-226.

BAILEY, K.M., and R. Ochsner. 1983. A Methodological Review of the Diary Studies: Windmill Tilting or Social Science? In K.M. Bailey, M.H. Long, and S. Peck (eds.), Second Language Acquisition Studies. Rowley, Mass: Newbury House, 188-198.

BATESON, G. 1972. Steps to An Ecology of Mind. New York: Ballantine.

BERTAUX, D. (ed.) 1981. ***Biography and Society:*** *The Life History Approach in the Social Sciences. Beverley Hills: Sage Publications.*

BROWN, C. 1983. ***The Distinguishing Characteristics of the Older Adult Language Learner.*** *Unpublished Doctoral Dissertation, University of California at Los Angeles.*

BROWN, C. 1985a. ***Requests for Specific Language Input:*** *Differences between Older and Younger Adult Language Learners. In S.L. Gass and C. Madden (eds.), Input in Second Language Acquisition. Rowley, Mass: Newbury House, 272-284.*

BROWN, C. 1985b. ***Two Windows on the Classroom World:*** *Diary Studies and Participant Observation Differences. In P. Larson, E.L. Judd, and D.S. Messerschmitt (eds.) ON TESOL '84: Brave New World for TESOL. Washington, D.C.: TESOL, 121-134.*

BROWN, R.H. 1977. ***A Poetic for Sociology:*** *Toward a Logic of Discovery for the Human Sciences. Cambridge: Cambridge University Press.*

BRUMFIT, C. and R. Mitchell (eds.). 1990. ***Research in the Language Classroom:*** *ELT Documents, 133. Modern English Publications and the British Council.*

CAVALCANTI, M. 1982. ***Using the Unorthodox, Unmeasurable Verbal Protocol Technique:*** *Qualitative Data in Foreign Language Reading Research. In S. Dingwall and S.J. Mann (eds.), Methods and Problems in Doing Applied Linguistic Research. Lancaster: University of Lancaster, 72-86.*

CAZDEN, C. 1988. ***Classroom Discourse.*** *Portsmouth, NH: Heinemann.*

CHAUDRON, C. 1983. ***Research on Metalinguistic Judgements:*** *A Review of Theory, Methods, and Results. Language Learning, 33(3):343-377.*

CHAUDRON, C. 1986. ***The Interaction of Quantitative and Qualitative Approaches to Research:*** *A View of the Second Language Classroom. TESOL Quarterly, 20(4):709-717.*

CHAUDRON, C. 1988. ***Second Language Classrooms:*** *Research on Teaching and Learning. Cambridge: Cambridge University Press.*

CHURCHLAND, P.M. 1990. ***Matter and Consciousness*** *(revised edition). Cambridge, Mass: The MIT Press.*

*COHEN, A.D. 1987. **Studying Learner Strategies:** How We Get the Information. In A. Wenden and J. Rubin (eds.), Learner Strategies in Language Learning. Englewood Cliffs, NJ: Prentice-Hall International, 31-42.*

*COHEN, A.D. and C. Hosenfeld. 1981. **Some Uses of Mentalistic Data in Second Language Research.** Language Learning, 31:285-313.*

*COOK, T. and C. Reichardt (eds.). 1979. **Qualitative and Quantitative Methods in Evaluation Research.** Beverly Hills: Sage Publications.*

*DANIELSON, D. 1981. **Views of Language Learning from an "Older Learner."** CATESOL Newsletter, 12 (January and March).*

*DENZIN, N.K. 1978. **The Research Act:** A Theoretical Introduction to Sociological Methods. (2nd ed.). New York: McGraw Hill.*

*DOBBERT, M.L. 1982. **Ethnographic Research:** Theory and Application to Modern Schools and Societies. New York: Praeger Publishers.*

*ELBOW, P. 1973. **Writing Without Teachers.** New York: Oxford University Press.*

*ELLIS, R. 1989a. **Classroom Learning Styles and Their Effect on Second Language Acquisition:** A Study of Two Learners. System, 17(2):249-262.*

*ELLIS, R. 1989b. **Sources of Intra-learner Variability in Language Use and their Relationship to Second Language Acquisition.** In Gass, S., C. Madden, D. Preston, and L. Selinker (eds.), Variation in Second Language Acquisition: Psycholinguistic Issues. Clevedon: Multilingual Matters.*

*ELLIS, R. 1990. **Learning Through Instruction:** The Study of Classroom Language Acquisition. Oxford: Blackwells.*

*ENRIGHT, L. 1981. **The Diary of a Classroom.** In J. Nixon (ed.), A Teacher's Guide to Action Research. London: Grant McIntyre, 37-51.*

*FAERCH, C. and G. Kasper (eds.), 1987. **Introspection in Second Language Research.** Clevedon: Multilingual Matters, Ltd.*

FIELDS, C.M. 1978. How Berlitz Taught Me Spanish Ràpidamente. The Chronicle of Higher Education. July 17,4-6, and July 24,7.

FRY, JOHN. 1988. Diary Studies in Classroom SLA Research: Problems and Prospects. Jalt Journal 9(2):158-167.

GAIES, S.J. 1983. The Investigation of Language Classroom Processes. TESOL Quarterly, 17(2):205-217.

GARDNER, R.C., P.C. Smythe, R. Clement, and I. Glicksman. 1976. Second Language Learning: A Social Psychological Perspective. Canadian Modern Language Review, 23: 198-213.

GEBHARD, J.G. 1984. Models of Supervision: Choices. TESOL Quarterly, 18(3):501-513.

GEERTZ, C.T. 1973. The Interpretation of Cultures. New York: Basic.

GEORGES, R.O. and M.O. Jones. 1980. People Studying People: The Human Element in Fieldwork. Berkeley: University of California Press.

GRANDCOLAS, B. 1986. Un Journal D'apprentissage Pour Préparar à L'enseignement. Paper presented at the TESOL France Conference, to be published in the conference proceedings.

GRANDCOLAS, B. and N. Soulé-Susbielles. 1986. The Analysis of the Foreign Language Classroom. Studies in Second Language Acquisition 8:293-308.

GREEN, J. and C. Wallat (eds.). 1981. Ethnography and Language in Educational Settings. Norwood, NJ: Ablex.

GREGG, K.R. 1984. Krashen's Monitor and Occam's Razor. Applied Linguistics, 5(2):79-100.

GROTJAHN, R. 1987. On the Methodological Basis of Introspective Methods. In C. Faerch and G. Kasper (eds.), 54-81.

GUBA, E. 1978. Toward a Methodology of Naturalistic Inquiry in Educational Evaluation. Monograph No. 8, Los Angeles: Center for the Study of Evaluation, University of California.

HAMMERSLEY, M. and P. Atkinson. 1983. Ethnography: Principles in Practice. London: Tavistock Publications.

*HATCH, E.M. 1978. **Second Language Acquisition: A Book of Readings.** Rowley, Mass: Newbury House.*

*HOWELL-RICHARDSON, C. and B. Parkinson. 1988. **Learner-diaries: Possibilities and Pitfalls.** In D. Grunwell (ed.), Language in Society: British Studies in Applied Linguistics, 3. London: CILT.*

*JOHNSON, J.M. 1975. **Doing Field Research.** New York: The Free Press.*

*JONES, R. 1977. **Psychological, Social, and Personal Factors in Second Language Acquisition.** Unpublished M.A. Thesis, English Department (ESL Section), University of California, Los Angeles.*

*JONES, R. 1980. **Social and Psychological Factors in Second Language Acquisition.** In C. Henning (ed.), Proceedings of the Los Angeles Second Language Research Forum. Los Angeles: University of California.*

*KAPLAN, A. 1964. **The Conduct of Inquiry: Methodology for Behavioral Science.** New York: Chander, Harper and Row.*

*KEMMIS, S. and J.A. Henry. 1989. **Action Research.** IATEFL Newsletter, 102:2-3.*

*KEMMIS, S. and R. McTaggart. 1982. **The Action Research Planner.** Victoria: Deakin University.*

*KRASHEN, S.D. 1983. **Newmark's "Ignorance Hypothesis" and Current Second Language Acquisition Theory.** In S. Gass and L.Selinker (eds.), Language Transfer in Language Learning. Rowley, Mass: Newbury House.*

*KREEFT-PEYTON, J. 1990. **Students and Teachers Writing Together: Perspectives on Journal Writing.** Alexandria, VA: TESOL.*

*KREEFT-PEYTON, J. and L. Reed. 1990. **Dialogue Journal Writing with Non-native English-Speakers: A Handbook for Teachers.** Alexandria, VA: TESOL.*

*KRIPPENDORFF, K. 1980. **Content Analysis: An Introduction to Its Methodology.** Beverly Hills: Sage Publications.*

*LABOV, W. 1972. **Some Principles of Linguistic Methodology.** Language in Society, 1:97-120.*

*LARSEN-FREEMAN, D. 1981. **Second Language Acquisition: Getting the Whole Picture.** In K.M. Bailey, M.H. Long, and S. Peck, (eds.), Second Language Acquisition Studies. Rowley, Mass: Newbury House, 3-22.*

*LINCOLN, Y.S. and E.G. Guba. 1985. **Naturalistic Inquiry.** Beverly Hills: Sage Publications.*

*LONG, M.H. 1980. **Inside the "Black Box": Methodological issues in Classroom Research on Language Learning.** Language Learning, 29(1):1-30; reprinted in H.W. Seliger and M.H. Long (eds.), 1983, Classroom Oriented Research in Second Language Acquisition. Rowley, Mass: Newbury House, 3-36.*

*MALINOWSKI, B. 1989. **A Diary in the Strict Sense of the Term.** Stanford: Stanford University Press (reprinted from the 1967 original, published by Routledge and Kegan Paul).*

*MANN, S.J. 1982. **Verbal Reports as Data: A Focus on Retrospection.** In S. Dingwall and S.J. Mann (eds.), Methods and Problems in Doing Applied Linguistic Research. Lancaster: University of Lancaster, 87-104.*

*MARSHALL, C. and G.B Rossman. 1989. **Designing Qualitative Research.** Beverly Hills: Sage Publications.*

*MATSUMOTO, K. 1987. **Diary Studies of Second Language Acquisition: A Critical Overview.** JALT Journal 9(1):17-34.*

*MATSUMOTO, K. 1989. **An Analysis of a Japanese ESL Learner's Diary: Factors Involved in the L2 Learning Process.** JALT Journal, 11(2):167-192.*

*MITCHELL, R. 1985. **Process Research in Second-language Classrooms.** Language Teaching, 18:330-352.*

*MOORE, T. 1977. **An Experimental Language Handicap (Personal Account).** Bulletin of the British Psychological Society, 30:107-110.*

*MORGAN, G. (ed.), 1983. **Beyond Method.** Beverly Hills: Sage Publications.*

*NUNAN, D. 1989. **Understanding Language Classrooms: A Guide For Teacher-initiated Action.** New York: Prentice-Hall International.*

NUNAN, D. 1990. Action Research in the Language Classroom. In J.C. Richards and D. Nunan (eds.), Second Language Teacher Education. Cambridge: Cambridge University Press, 62-81.

OCHSNER, R. 1979. A Poetic of Second Language Acquisition. Language Learning, 29:53-80.

OLLER, J. and K. Perkins. 1978. Intelligence and Language Proficiency as Sources of Variance in Self-reported Affective Variables. Language Learning, 28:85-97.

OXFORD ENGLISH DICTIONARY (Compact Edition). 1971. Glasgow: Oxford University Press, 718-719.

PARKINSON, B. and C. Howell-Richardson, 1990. Learner Diaries. In C. Brumfit and R. Mitchell (eds.).

PATTON, M.Q. 1980. Qualitative Evaluation Methods. Beverly Hills: Sage Publications.

PLUMMER, K. 1983. Documents of Life: An Introduction to the Problems and Literature of a Humanistic Method. London: George Allen and Unwin.

POPKIN, D. 1985. Dialogue Journals: A Way to Personalize Communication in a Foreign Language. Foreign Language Annals, 18(2):153-156.

REASON, P. (ed.). 1988. Human Inquiry in Action: Developments in New Paradigm Research. Beverly Hills: Sage Publications.

REASON, R. and P. Hawkins. 1988. Storytelling as Inquiry. In P. Reason (ed.), 79-101.

RIVERS, W.M. 1979. Learning a Sixth Language: An Adult Learner's Daily Diary. Canadian Modern Language Review, 36:1(67-82); reprinted in Rivers, W.M. 1983. Communicating Naturally in a Second Language: Theory and Practice in Language Teaching. Cambridge: Cambridge University Press, 169-188.

RUBIN, J. and R. Henze. 1981. The Foreign Language Requirement: A Suggestion to Enhance its Educational Role in Teacher Training. TESOL Newsletter, 15(1):17, 19, 24.

SAVILLE-TROIKE, M. 1982. The Ethnography of Communication: An Introduction. Baltimore: University Park Press.

*SCHMIDT, R.W., and S.N. Frota. 1986. **Developing Basic Conversational Ability in a Second Language: A Case Study of an Adult Learner of Portuguese.** In R.R. day (ed.), Talking to Learn: Conversation in Second Language Acquisition. Rowley, Mass: Newbury House, 237-326.*

*SCHUMANN, F.E. 1980. **Diary of a Language Learner: A Further Analysis.** In R. Scarcella and S. Krashen (eds.), Research in Second Language Acquisition: Selected Papers of the Los Angeles Second Language Research Forum. Rowley, Mass: Newbury House, 51-57.*

*SCHUMANN, F.E. and J.H. Schumann. 1977. **Diary of a Language Learner: An Introspective Study of Second Language Learning.** In H.D. Brown, R.H. Crymes, and C.A. Yorio (eds.), On TESOL '77: Teaching and Learning English as a Second Language -- Trends in Research and Practice. Washington, D.C.: TESOL, 241-249.*

*SCHUMANN, J.H. 1983. **Art and Science in Second Language Acquisition Research.** In M.A. Clarke and J. Handscombe (eds.), On TESOL '82: Pacific Perspectives on Language Learning and Teaching. Washington, D.C.: TESOL, 107-124.*

*SELIGER, H.W. 1983. **The Language Learner as Linguist: Of Metaphors and Realities.** Applied Linguistics 4(3):179-191.*

*SHARWOOD-SMITH, M. 1981. **Consciousness-raising and the Second Language Learner.** Applied Linguistics, 2(2):159-168.*

*SMITH, J.K. and L. Heshusius. 1986. **Closing Down the Conversation: The End of the Quantitative-Qualitative Debate among Educational Inquirers.** Educational Researchers, 15(1):4-12.*

*SPACK, R. and C. Sadow. 1983. **Student-teacher Working Journals in ESL Freshman Composition.** TESOL Quarterly, 17(4):575-593.*

*SPRADLEY, J.P. 1979. **The Ethnographic Interview.** London: Holt, Rinehart and Winston.*

*SPRADLEY, J.P. 1980. **Participant Observation.** London: Holt, Rinehart and Winston.*

*STATON, J. 1981. **Literacy as an Interactive Process.** The Linguistic Reporter, 24(2):1-5.*

*TRUEBA, H.J., G.P. Guthrie, and H.P. Au (eds.). 1981. **Culture and the Bilingual Classroom: Studies in Classroom Ethnography.** Rowley, Mass: Newbury House.*

van LIER, L. 1984. *Discourse Analysis and Classroom Research: A Methodological Perspective. International Journal of the Sociology of Language, 49:111-133.*

van LIER, L. 1988. *The Classroom and the Language Learner: Ethnography and Second Language Classroom Research. London: Longman, Inc.*

van LIER, L. 1989. *Classroom Research in Second Language Acquisition. Annual Review of Applied Linguistics, 10:173-186.*

van LIER, L. 1990. *Ethnography: Bandaid, Bandwagon, or Contraband? In C. Brumfit and R. Mitchell (eds.), 33-53.*

WATSON, L.C. 1976. *Understanding a Life History as a Subjective Document: Hermeneutical and Phenomenological Perspectives. Ethos, 4(1):95-131.*

WATSON-GEGEO, K. 1988. *Ethnography in ESL: Defining the Essentials. TESOL Quarterly, 22(4):575-592.*

YIN, R.K. 1984. *Case Study Research: Design and Methods. Beverly Hills: Sage Publications.*

PERSONALITY AND SECOND-LANGUAGE LEARNING: THEORY, RESEARCH AND PRACTICE

Roger Griffiths

Abstract

Despite the importance of differences between language learners being frequently asserted in the second-language literature, investigations of the relationship between personality and learning have largely failed to produce significant findings. Consequently, reviewers have tended to write off personality variables from the L2 research agenda. The claim made in this paper is that this is unjustified, and that it results from giving unwarranted credibility to studies whose basic assumptions are ill-founded. Theoretically sound and research based hypotheses derived from the psychological literature (particularly the work of Eysenck) are, however, proposed as alternatives to global correlational ones. In reviewing specialist research in the field, areas in which personality is seen to be of potential importance in an L2 context include: methods; specific methodology; task-based learning; pairwork/groupwork; praise/reinforcement; range of stimuli; and testing. Cross-cultural research is demonstrated to be of particular interest in Asia where a number of distinctive personality and mental ability profiles have been observed. Preliminary findings from on-going research are reported. Research findings on the personality of teachers are also discussed. It is concluded that by working within established theory and using validated instruments, research on personality in an L2 context has demonstrable potential to inform practice.

Introduction

Despite resurgent interest in individual differences of second language learners (e.g. Skehan, 1989; de Jong & Stevenson, 1990) the study of personality as a variable in L2 learning is in a state of serious decline; it has in fact largely been written off the research agenda (as reflected in the publication of major journal articles).

However there are grounds for maintaining that the adoption of an alternative research perspective is long overdue. Four arguments are consequently proposed in this paper.

1) The writing off of personality variables in the context of L2 research is unjustified, and results from giving credibility to studies based on ill-conceived assumptions.

2) A sufficiently detailed and elaborated theory of personality exists and could be profitably built on.

3) Mainstream psychological research indicates variables and interactions which might fruitfully be investigated in second-language classrooms, within the theoretical framework referred to above.

4) Research carried out in this way is likely to be particularly relevant in cross-cultural studies due to the reported distinctiveness of national/racial personality profiles and mental ability profiles (especially in Asia).

Present Position Accorded to the Importance of Personality Variables in Second Language Research

Whereas the importance of affective variables in second-language learning is frequently acknowledged in the L2 literature (e.g. Brown, 1987, p. 111; Bley-Vroman, 1989, p. 49; but see Long, 1990, p. 275 for contrary view), variables specifically within the realm of personality are currently accorded little importance in research reviews. This is due to the fact that studies in which the role of personality variables have been investigated in relation to language learning have failed to produce consistently significant findings, e.g. Swain and Burnaby (1976) Naiman, Frolich, Stern, and Tadasco (1978), Genessee and Hamayan (1980).

The hypothesis which has received the most attention (actually almost all the attention) is that relating extraversion to language learning proficiency; the relationship is however very far from being definitively established, and reviewers have consequently arrived at generally pessimistic conclusions as to the importance of personality variables in SLA, and even questioned the validity of the constructs investigated. For example, in one of the most recent reviews, Skehan (1989),

maintains that L2 research should concern itself with improving definitions in the personality domain. He states:

> "There are grounds for questioning the desirability of adopting, wholesale, a construct from a feeder discipline, psychology, rather than subjecting the construct to further analysis to relate it specifically to language learning." (p. 105)

In like vein, McDonough (1986, p. 139) reports that L2 researchers have not always been satisfied with the validity of personality scales developed for psychological studies. Ellis (1985, p. 120) similarly observes that L2 researchers have often preferred to develop their own batteries of personality traits "which intuitively strike them as important".

Summarising findings from L2 studies undertaken in this area, Ellis (1985) also notes that "In general the available research does not show a clearly defined effect of personality on SLA" (p. 121). He further states that the major difficulty of personality research in a second language context is that of identification and measurement:

> "At the moment, a failure to find an expected relationship (e.g. between extroversion and proficiency) may be because the test used to measure the personality trait lacks validity." (p. 122)

However, it can be maintained that Ellis is no more right on this issue than is Skehan in doubting the wisdom of directly adopting constructs from psychology. It is, in fact, the central argument of this paper that only by working within the wider theoretical framework of personality theory as elaborated in the psychological literature, and only by making use of the tests which arise from that theory, will real progress be made in investigating personality within the context of second language learning.

Insofar, therefore, as L2 reviewers have got it wrong, it is clearly necessary to demonstrate how and why they did so. They do, after all, draw their conclusions from the results of extant studies. However, if the fundamental assumptions of studies, rather than their findings, are examined, an alternative explanation becomes possible. The study by Busch (1982), mentioned above and described by Brown (1987, p. 109) as "the most comprehensive study to date on extroversion", furnishes a representative example.

Busch begins her paper by stating:

> "The assumption that there is a relationship between extraversion and proficiency in a foreign language is widely held by teachers, researchers, and students of second languages." (p. 109)

She then describes the evidence which she supposes supports this assumption (it does after all provide the raison d'etre of her study).

The basis for including students in the above generalisation is given first. She notes that "31% of the students who were considered to be good language learners in the study by Naiman, Frohlich, and Stern (1975) stated that extraversion was helpful in acquiring oral skills". However, not only is the reference only (and merely) to oral skills but she might have observed (but did not) that 69% of the students did not state that extraversion was helpful.

Secondly, to justify the claim that the assumption is widely held by researchers, she cites comment by Rubin (1975). Rubin, however, merely states that good language learners have a strong desire to communicate (in Busch, 1982, p. 109). The opinions of two other researchers who are supposed (somewhat tangentially) to support the assumption, only appear in unpublished studies, and Busch actually quotes Brown (1973) as questioning the view that introverts are qualitatively less proficient than extraverts. In all, this is not convincing evidence of the assumption being widely held by researchers (even in L2).

No data are given to support the assumption attributed to teachers.

Although Busch clearly fails to justify her claim that there is widespread support for the assumption relating extraversion to language learning, she maintains that "psychologist have written volumes on the subject of introversion-extraversion" (p. 110). This is indeed true and she might have added that in reading "extraverts are more proficient in English" (Busch's hypothesis, 1982, p. 109). Indeed, the fatal flaw in this study, and in others which have looked at global language proficiency and personality, results from the postulating of naive relationships. It is quite simply the case that thus far the hypotheses investigated in L2 studies of personality are neither logically predicted from personality theory nor would they be anticipated from a reading of the relevant experimental literature. (In fact, insofar as general findings are available on learning and personality, that of Wankowski [1978] from an extensive study of British university students is that,

"Generally speaking, it seems that in both general and department populations, it is the neurotic and extravert students who tend to be less successful.." [p. 43-44]). Consequently, the fact that researchers have not found relationships cannot fairly be used (as it has been) to dismiss personality variables from the L2 research agenda; nor can highly validated psychometric instruments be held accountable for the failure.

However, if it is accepted that what has gone before has not gone very far and should arguably never have started out, then a number of things need to be accounted for. Firstly, there is researchers' evident interest in the area (as demonstrated by the studies undertaken). Secondly, there are findings from L1 research (e.g. Blease, 1986) that teachers regard personality variables as of considerable importance in learning. Thirdly, a survey of teachers ratings of the importance of psychological variables in language learning, shows that L2 teachers have at least as high a regard for personality factors as do content teachers.

(The survey was conducted with university and language school teachers in 3 countries; Japan, England and Oman. The aim of the survey, which was regarded as a very simple and initial probe into teachers' opinions, was to investigate whether these variables were seen as important by actual classroom practitioners. Responses to a question regarding the importance of 3 psychological variables [intelligence, personality and memory] in successful classroom language learning were recorded on a 5-point Likert scale anchored by very important (5) and not important (1). All three variables are observed to be highly regarded, with only minimal differences being observed between them. As the survey is recognised as an extremely simple one, further analysis of these findings is not justified, but the figures on "personality" do show, as hoped, that the area might have more potential than a reading of the extant research findings suggests).

It is, I suggest, possible to account for the perceived importance of personality variables, by looking, not at some understandably elusive global correlation between extraversion and language proficiency, but by exploring interactions and micro-areas where both theory and experimental evidence indicate the possibility of observing predicted relationships in language classrooms. However, before that can be done, as there are a number of alternative personality theories to choose from, justification for using a particular theory needs to be established and the theory itself needs to be briefly described.

Table 1

Psychological Variables Survey Findings

Country	N	Personality		Intelligence		Memory	
Japan	21	4.2	1.01	3.9	0.81	3.9	0.81
Oman	20	3.8	0.95	4.2	0.76	3.8	0.89
England	57	4.0	1.01	3.7	0.81	4.3	0.86
Total	98	4.0	0.99	3.9	0.80	4.1	0.86

(Max. = 5)

Eysenck's Theory of Personality

Psychology books on personality are usually arranged eponymously (e.g. Burger, 1986, has sections on Freud, Erikson, Eysenck, Kelly, etc.). Such compilations present tantalizing glimpses of research-based wealth amongst data-less poverty dressed in elaborate metaphors and expansive reflection. Consequently, if the need is for a comprehensive theory based on empirical studies which gives rise to testable hypotheses, preferrably having a history of experimental verification, then the choice narrows considerably. If, in addition, there is a need for a reliable and verified personality test which is derived from the theory, the choice narrows even further. In fact, almost axiomatically, it leads to the choice being made from "trait" theories of personality.

The model described in this paper is that of H.J. Eysenck (e.g. Eysenck, 1967, 1970; H.J. Eysenck & S.B.G. Eysenck, 1975; H.J. Eysenck & M.W. Eysenck, 1985). Eysenck (1970) defines **personality** as:

> "A more or less stable and enduring organization of a person's character, temperament, intellect, and physique, which determines his unique adjustment to the environment. **Character** denotes a person's more or less stable and enduring system of conative behaviour (**will**); **temperament**, his more or less stable and enduring system of affective behaviour (**emotion**); **intellect**, his more or less stable and enduring system of cognitive behaviour (**intelligence**); **physique**, his more or less stable and enduring system of bodily configuration and neuroendocrine endowment." (p. 2).

Eysenck's model is preferred over possible alternatives (e.g. Cattell, 1957) as it incorporates a well-validated taxonomy within an explicit causal theoretical framework. Also, not only has the theory itself generated a vast amount of (largely supportive) research, but the psychometric instrument used to measure the major factors (the Eysenck Personality questionnaire, [EPQ], Eysenck & Eysenck, 1975) has been standardized in some 35 different countries (several of these being in Asia).

Although Eysenck's model addresses both taxonomy and causality, clearly, taxonomy is primary: if personality could not be reliably measured on standardized instruments then causal theories would merely be verbal effusions on the elusive and ethereal. the notion of a hierarchic structure of behaviour (specific response - habitual response - trait - type) is central to Eysenck's model; it culminates in defining the type-concepts of introversion-extraversion (I-E), emotionality (otherwise described as neuroticism ([N]-stability), and tough-mindedness (otherwise known as psychoticism [P]-normality/impulse control). These major dimensions, which are represented in almost all large-scale studies and nearly all theoretical formulations (e.g. Cattell, 1957; Digman, 1989), are represented by continua, the extremes of which can be described through idealized types:

> "Extraverts are sociable, like parties, have many friends and need excitement; they are sensation-seekers and risk-takers, like practical jokes and are lively and active. Conversely introverts are quiet, prefer reading to meeting people, have few but close friends and usually avoid excitement. Subjects who score high on emotionality tend to be worriers, often anxious, moody and sometimes depressed; they over-react to stress, finding it difficult to calm down afterwards. Stable individuals on the other hand, are usually calm, even-tempered and unworried; they are slow to respond emotionally and recover their equipoise quickly after arousal. Tough-minded people are characterized by aggressive, hostile behaviour; they seem cold emotionally, lack empathy and are insensitive to the feelings of other people as well as their own; they are impulsive and egocentric but often also original and creative. They tend to be unconventional and appear to like odd, unusual people and things" (Eysenck & Chan, 1982, p. 154)

Test-retest reliabilities of the E and N scales in the very extensive British Standardization are reported to be .89 and .86 respectively.

The question of the validity of the EPQ (i.e. does the test actually measures what it is intended to measure), is extensively commented on by Eysenck and against which the test can be evaluated (obviously, the existence of such a criterion would make the test unnecessary); the answer lies in looking at the various ways in which validity can be approached. Content and face validity are clearly of little relevance as the appropriacy of items need to be assessed using statistical techniques, and a priori selection does not guarantee that items actually load on factors.

Eysenck and Eysenck use the term "consensual" (p. 77) validity to describe questionnaire validation through comparison with external ratings of informed observers. They report satisfactory evidence in respect of P, E and N, in this area. They also note that predictive validity of the EPQ is shown by correlations between ratings at one age and questionnaire results at another. Support for the validity of the major dimensions is also attained through criterion analysis. It is, in fact, a particular feature of Eysenck's conceptualisation that he insists that a criterion group anchor each dimension (hence Eysenck's retention of the factor names of neuroticism and psychoticism, instead of the less emotionally loaded labels of "emotionality" and "tough-mindedness"). According to Eysenck and Eysenck (1985) the question of construct validity (the extent to which a test may be considered to measure a particular theoretical construct), "should [ideally] involve a much more abstract type of theory, making possible far more complex and surprising predictions than would be possible on... [a] simple descriptive model" (p. 81). Such a theory is developed in the account of causality proposed by Eysenck.

Eysenck accounts for introverted and extraverted behaviour by reference to the concept of cortical arousal. He proposes that extraverts are characterized by under-arousal, introverts by high arousal. Extraverts are consequently driven to increase arousal through sensation-seeking etc., and introverts, who are chronically more cortically aroused, attempt to avoid strongly arousing stimuli. Eysenck and Eysenck (1985, p. 208) acknowledge that arousal theory provides a somewhat imprecise and oversimplified perspective, but an extremely large research literature derived from both physiological and psychological testing has produced results which have been essentially confirmatory; relationships between P,E and N and a host of behaviours such as accident proneness, anti-social behaviour, criminality, and smoking, are well documented (e.g. Eysenck, 1976). There is also abundant and convincing evidence for a genetic account of personality derived from twin studies (e.g. Floderus-Myrhed, Pederson, & Rasmuson, 1980; Rose, Koshenvuo,

Kaprio, Sarna, & Langinvainio, 1988), and evidence of physiological differences between introverts and extraverts, e.g. in temporal lobe blood flow distribution (Stenberg, Risberg, Warkentin, & Rosen, 1990).

While it will be recognised that the above account of Eysenck's theory is extremely condensed, it is hopefully sufficient to demonstrate its potential application in L2 research. Presuming this to be the case, it therefore becomes necessary to report general educational research findings and to relate these to areas of L2 research concern. This is done in the next section.

Identifying Meaningful Hypotheses

To the extent that the evidence showing that personality features interest with learning in meaningful ways supports teachers' beliefs (as expressed in the survey reported earlier) that personality is important in learning, it becomes the responsibility of L2 researcher to identify the domains in which such factors operate; this can only be done through consulting the psychological literature and relating both theory and findings to the L2 learning situation. However, as stated previously, as significant global correlations have only very infrequently been observed, this necessitates an analysis of specific interactions and micro-areas; those in which relevant research findings exist include the following:

1) Methods
   discovery/receptive learning

2) Specific methodology
   position of rules

3) Pairwork/groupwork

4) Tasks

5) Praise/reinforcement

6) Range and volume of stimuli
   a) boredom
   b) noise

7) Testing

Given the particular focus of this paper, and considering that studies relating to all of the above have been reviewed elsewhere (Griffiths, 1991), clearly not all of these areas need to be covered in detail. However, selected examples of relevant research will be described.

Methods

The conclusion that different methods produce the same results and tend to give support only to the null hypothesis is regarded by Eysenck (e.g. 1978, p. 145) as artifactual. He claims that introducing personality into the picture "often serves to demonstrate quite clearly that two methods apparently equal in effectiveness, differ sharply, one being much better for introverts, the other for extraverts" (1984, p. 185).

The truth of this claim is demonstrated by findings from a series of studies undertaken by Leith (e.g. 1969, also reported, 1974), one of which investigated differences in responses to "reception learning" (i.e. standard deductive teaching of principles by direct instruction) and "discovery learning" (i.e. the inductive method in which students are asked to find out principles and results for themselves). Leith investigated the hypothesis that "the greater readiness of extraverts to become bored by routines but likely to respond to stimulus variation, and of introverts to be disturbed by changes of set but able to maintain attentiveness to a highly prompted task, would result in a methods by personality interaction" (cited in Eysenck, 1978, p. 145).

A carefully prepared course (in genetics) was delivered to 211 students in randomly assembled treatment groups. One group learned from a program prepared in the form of direct instruction (reception learning), while the other learned from a discovery program. The discovery program was organised so that a complex whole was given first and then analysed, while the reception learning version built up the complex whole step by step. Both programs contained exactly the same content and the same examples were covered in each version. The difference was essentially one of induction or deduction. A post-test on the material taught and relevant transfer items was given one week after the completion of the course, and a retention test was given after a further four weeks.

Both sets of scores showed significant treatment x extraversion interactions ($p < .05$ & $p < .01$). Findings from the experiment are given below in both tabled and diagramatic forms.

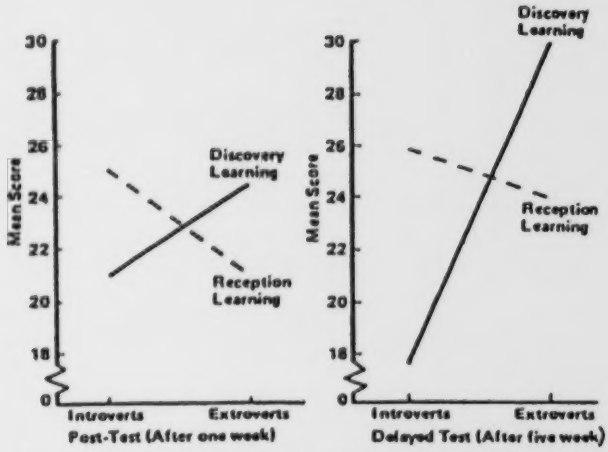
Table 2

MEAN POST TEST SCORES OF INTROVERTS AND EXTRAVERTS GIVEN RECEPTION LEARNING AND GUIDED DISCOVERY PROGRAMMES

	Teaching Strategy	
Personality Type	Reception	Discovery
Extraverts	21.37	24.28
Introverts	24.67	21.03

MEAN RETENTION TEST SCORES OF INTROVERTS AND EXTRAVERTS GIVEN RECEPTION LEARNING AND GUIDED DISCOVERY PROGRAMMES

	Teaching Strategy	
Personality Type	Reception	Discovery
Extraverts	24.44	29.63
Introverts	25.72	17.35



Effects of discovery and reception learning in introverts and extraverts (from Leith, 1974).

FIG. 1 FIG 2

Leith concludes: "The notable point about [this] experiment is that, unless differences in personality had been included in the design, the methods of presenting learning tasks would have appeared to give the same results" (1969, p. 108).

Similar findings have resulted from other investigations, for example, Amaria and Leith (1969), and Leith and Wisdom (1970).

Insofar as it is possible to compare these findings (taken from content teaching) to language teaching (especially in regard to communicative language teaching which makes use of the principle of inductive learning, compared with deductive methods which emphasize the study of grammatical rules), the finding that discovery/inductive methods favour extraverts while reception/deduction favours introverts, must surely excite interest. However interactions of method and personality (of both students and teachers) appear not to have been investigated in this context.

Pairwork/Groupwork

The current emphasis on pair-work in L2 teaching and the claim that small-group work involving two-way tasks promotes learning through the negotiation of input (e.g. Long & Porter, 1985, p. 224) makes this an obvious area of investigation in terms of personality variables.

Findings from personality research, do in fact show that the dimensions of both I-E and N are relevant to performance in groups and pairs. In a study by Leith (1974), for example, pairs were formed on the basis of the personality variable of anxiety/neuroticism, pairs having either similar scores on this variable, or opposite (i.e. one anxious, the other stable). The main results are best shown in the form of a table.

Table 4

*Comparisons of achievements and behaviour of same and different anxiety level pairs**

(Heterogeneous ability pairs in brackets, homogeneous ability pairs without brackets)

Opposite anxiety pairs	Achieved	74% (32%)	more on the post-test than same anxiety pairs
Opposite anxiety pairs	Achieved	98% (113%)	more on the transfer-test Than same anxiety pairs
Opposite anxiety pairs	Spent	59% (36%)	more time in showing solidarity, raising other's status, giving help and rewarding than same anxiety pairs
Opposite anxiety pairs	Spent	121% (132%)	more time asking for orientation, information, confirmation, than same anxiety pairs
Opposite anxiety pairs	Spent	11% (20%)	less time in disagreeing passively rejecting, withholding help than same anxiety pairs

* From Leith, 1974.

Footnote

The distinction between methods in this area of research is seen in terms of the degree of guidance given to the learner. Ambiguities in the use of the term "method" are acknowledged, but, in general terms, discovery learning is considered to involve the least amount of teacher guidance, and direct instruction the most.

Remarkable improvements over the "'same' pairings are shown in 'unlike' pairings". Opposite anxiety pairs in the transfer test and, of 100% superiority over the same pairs in the transfer test and, of particular importance in the light of L2 input-negotiation findings (e.g. Long & Porter, 1985), they also spent 121% - 132% more time asking for orientation, information, and confirmation.

While it is obviously necessary to replicate such studies in an L2 context, it should certainly be the case that, as Eysenck claims, "These results open up fascinating vistas for both research and educational practice" (1978, p. 151). This must be particularly true in view of the importance ascribed to pair work in communicative language teaching.

Tasks

Evidence such as that described above does at least suggest the possibility that personality is important in, for example, determining individual task-type preference, determining response to tasks of varying levels of difficulty, even general disposition to a task-based approach. Specific hypotheses related to personality variables therefore appear to be worth exploring in relation to the use of tasks in language classrooms.

Praise and Reinforcement

The standard teach-training edict to praise success, is not necessarily supported by experimental evidence, (e.g. McCullers, Fabes, & Moran, 1987) clearly shows that giving rewards (toys to young children in this case) can have adverse effects on immediate performance. Other research also shows that responses are often related to personality dimensions. Nagpal and Gupta (1979), for example, found clear evidence that individuals differ in their susceptibility to reinforcement: "Extraverts condition more readily with the rewarding reinforcers while introverts condition more readily with the punishing reinforcers..." (p. 475)

The need to bring findings such as these to the attention of language teachers, and to conduct specific L2 classroom research in this area is clearly obvious.

Range and Volume of Stimuli

Boredom

The pervasive finding of experimental work on boredom (e.g. Wankowski, 1978, p. 47) is that it is very much in the eye of the beholder. This fact must be clearly recognised in the language classroom where teachers might overreact to the need for activity change voiced by highly vocal extraverts while introverts might be quietly contented.

Noise

Experimental evidence also suggests that the question of classroom noise levels might best be approached through looking at its effects on different personality types. Dornic and Ekehammer (1990), for example, in a study with 215 Swedish university students, found a highly significant negative correlation between noise sensitivity and extraversion, while Campbell and Hawley (1982) found higher extraversion scores for students who preferred a noisy library reading room to those who preferred a quieter room.

To the extent that classroom noise levels differ (and clearly they do both within and between classes), it appears that performance by introverts and extraverts is also likely to differ. The L2 researcher could therefore be involved in looking at the effect of noise level on the performance of students of differing personalities while involved in different activities.

Testing

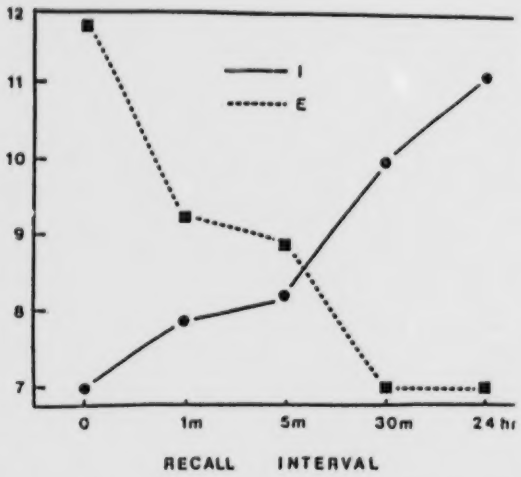
There are numerous branches of research on personality differences which are of direct relevance to assessment and testing, and the field clearly needs to be approached from both major dimensions of personality, I-E and N.

Findings from research on memory are clearly relevant to testing, and indicate the importance of personality in this area. At the end of a chapter devoted to describing differences in learning and memory between introverts and extraverts, M.W. Eysenck (1977, p. 217), for example, concludes "The experimental evidence indicates that there are relatively consistent differences in extraverts show more

rapid learning than introverts on difficult tasks, such as those involving response competition [see Eysenck 1977, p. 184]; extraverts tend to recall better than introverts at short retention intervals; and extraverts retrieve information faster than introverts from episodic and semantic memory".

Findings of Howarth and H.J. Eysenck (1968) show this relationship in a particularly striking fashion.

Figure 3



scores of extraverts and introverts at the recall interval stated.
Maximum recall 14.
Each point is the mean score of 11 *Ss*.
From Howarth and Eysenck, 1968.

H.J. Eysenck (1978) has attempted to relate such findings to an educational context; he observes, "Clearly introverts remember better in the long run, while consolidation is still proceeding; if quizzed during this time, they may give the impression of not having paid attention. Extraverts may shine in the short period after learning, but will disappoint in the long run..." (p. 169)

The timing of testing is therefore seen to be of importance; a quiz given straight after a presentation being likely to favour extraverts while, at a later date, introverts are likely to show gain scores.

A number of other hypotheses related to general educational testing have been identified (Griffiths, 1991), and several of these are currently being related specifically to the second-language context, and subsequently investigated. Berry (forthcoming), for example, used the 86-item EPQ previously validated in Japan (Iwawaki, Eysenck &Eysenck, 1980) to identify groups of extreme introverts and extraverts in a sample of 96 junior college girls, and results of an English language achievement test/subtests were related to I-E scores. As predicted, no significant differences were obtained when overall test scores were analysed, but extraverts were shown to score significantly lower than introverts ($p = .03$) on a phrasal-level gap-fill dictation subtest. This result was also predicted from the personality literature which indicates extraverts to be more impulsive and less accurate on this type of test.

These are of course preliminary findings and little can be made of them until they are replicated, but the fact that they emerged from a study based on a test designed to assess achievement on a specific course (while in the specialist literature careful control of variables is often required before main effects are discerned, e.g. Revelle, Humphreys, Simon, and Gilliland, 1980) does add to their conviction.

In general, while it would be simplistic to expect to find any gross differences in language test scores between personality types it does however seem that particular types of test (and/or item types), modes of presentation, testing context, processing skills being assessed, etc., will be differentially responded to by extraverts-introverts/ highN-lowN learners.

Another area related to personality from which hypotheses are currently being derived is that of mental ability, where the visuo-spatial/verbal distinction is attracting contemporary attention.

Mental Ability Profiles

As noted earlier, Eysenck's definition of personality includes a cognitive dimension (intelligence) and here too, a number of Asian groups are observed to

have distinctive profiles, especially in the balance of verbal and non-nonverbal abilities. This is of particular interest as it is an area in which hypotheses are currently being formulated in the L2 literature.

It is suggested, for example, that case-study findings of exceptional language learning ability being observed in subjects with relatively weak visuo-spatial abilities (e.g. as reported in Novoa, Fein and Obler, 1988, p. 301) support the view of Schneiderman and Desmarais (1988) that mild to severe disabilities in visuo-spatial functions may be a frequent concomitant of L2 aptitude. They propose that talented language learners are "less left-lateralized for language than individuals who are less flexible and consequently less talented for second-language learning" (p. 116); as, in neurological terms, verbal abilities are, broadly, localised in the left hemisphere and visuo-spatial abilities in the right (Lynn, 1987, p. 814), then a trade-off of abilities appears possible, and the proposal has prima facie support. However, as the tests used in such studies (e.g. analogy tasks in matrix form) are generally regarded as good measures of general ability (e.g. Cronbach, 1990, p. 231-232), and as this has been shown to have a positive, if not very strong, relationship with successful L2 learning (e.g. Wesche, Edwards, & Wells, 1982) then the hypothesis must still be regarded as a very preliminary one. It is however supported by, for example, Smith (1964, cited in Lynn, 1987, p. 839-840) who, inter alia, found a negative correlation between spatial ability and examination performance in German by British school children. In general, however, the position is that while the usefulness of the mental ability verbal/visuo-spatial distinction has not been firmly established in an L2 context, the available evidence appears to make it worth further investigation.

It is therefore clear that hypotheses relevant to language learning can be derived from the psychological literature. Consequently, if personality can be assumed to be important in classroom language learning, the information on individual differences at all levels becomes worthy of consideration. Differences between races and nations (specifically referring to Asian nations) are therefore considered in the next section.

Asian Personality Profiles

In L2 research, as in science generally, we must look at regularities in nature as well as at the unique individual. We can of course do this at a number of levels all of which (in the case of personality and language learning) might be regarded as of

value. At the most general level, we are likely to derive some information from data at both the national and racial levels.

Similarities in personality structure between groups have been extensively observed (Barrett & Eysenck, 1984), and a good deal of standardization data is available for some 35 countries including Singapore, Japan, Hong Kong, and India. In order to facilitate cross-cultural comparisons, Barrett and Eysenck, have presented data on transformed scales (which mean that "the figures are directly comparable to one another" [p. 617]) from 25 countries. Those of regional interest are given in Table 5.

Table 5

Transformed Scale Means for Cross-Cultural Comparison

Males

Country	Sample	E	N	P	L
Australia	336	18.83	14.88	8.41	7.39
China	500	14.49	13.85	7.62	19.49
Hong Kong	270	17.39	14.24	8.36	13.62
India	509	23.14	14.67	8.41	17.71
Japan	719	16.28	16.13	5.32	9.01
Singapore	493	18.46	11.61	4.74	15.77
U.K.	600	17.97	12.77	4.79	10.89

Females

Country	Sample	E	N	P	L
Australia	318	19.79	16.08	5.50	7.76
China	500	13.01	15.14	5.95	21.33
Hong Kong	462	16.06	14.97	5.74	15.12
India	472	22.45	17.84	7.92	19.04
Japan	599	16.71	17.43	4.28	10.23
Singapore	501	16.38	14.43	3.97	16.87
U.K.	598	18.09	17.17	2.89	13.33

Each scale mean is presented as though derived from a <30>-item scale

(From Barrett & Eysenck, 1984, p. 618)

As is apparent, both similarities and differences between national populations emerges from the analysis. In line with the earlier discussion of personality variables which appear to be relevant to language teaching/learning and, as demonstrated by Berry (forthcoming), language testing bias, differences between scores on E and N are of particular interest. The very high E scores in China and Japan. Similarly, the high N scores of the Japanese (especially males) contrast with most of the other data.

The Japanese are in fact observed to have higher scores on introversion (lowE) and neuroticism (emotionality), than citizens of most other countries (apart from China). Considerable support for this observation, originally based on a series of studies by Iwawaki et al. (e.g. 1980), has recently been obtained in a large scale study (n = 609 male and female college and university students) using a reduced scale 86-item EPQ, by Griffiths and Berry (in preparation) and a smaller study (n - 181 male university students) by Griffiths, and Sheen (in preparation, a).

Another confirmatory finding of on-going research (Griffiths & Sheen, in preparation, a) is that of the high visuo-spatial ability of the Japanese (Lynn, 1987). Having rejected the GEFT (Witkin, Oltman, Raskin & Karp, 1971) as a measure of field independence/independence (Griffiths & Sheen, in preparation, b), but viewing it as a measure of visuo-spatial ability and an excellent measure of fluid intelligence (following Cronbach, 1984, p. 265), it was administered to 175 male students enrolled at a middle ranking Japanese private university (and therefore probably of no more than high average IQ) and 63 female college students. Scores considerably higher than USA norms and also even higher than those previously reported for 112 Japanese subjects in a study of 816 foreign students in the English Language Institute at Brigham Young University-Hawaii (e.g. Hansen-Strain & Strain..., p. 260) were observed. Table 6 shows relevant scores from both studies (Note 1: only scores obtained with n >40 are reported. Note 2: USA norms for college students; males=12.0, s.d. 4.1; females=10.8, s.d. 4.2; Witkin et al., 1971, p. 28).

Table 6

GEFT Scores in Two Asian Studies

Group	N	Mean	s.d.
Samoa	83	9.9	4.4
Tonga	139	8.4	5.5
Micronesia	63	6.8	4.0
Philippines	42	6.5	4.6
Hong Kong	194	13.7	4.3
Chinese	55	13.3	3.6
Korea	103	10.4	4.7
Japan	112	14.1	4.5
Japan (Males)		15.8	
Japan (Females)		13.3	

(Hansen-Strain & Strain, 19... p. 260)

Japan (Males)	175	16.0	3.0
Japan (Females)	63	15.2	3.0

(max. score = 18)

(Griffiths & Sheen, in preparation, a)

Insofar as these findings can be thought of as representative comparable samples of the nations in question (and Hansen-Strain p.c. ....., but Werner, [1979, cited in Royce, 1988, p. 159], indicates higher Filipino spatial ability than shown here), results of testing on the GEFT are clearly seen to vary considerably between nations, thus confirming Jenson and Reynolds' (1982; 423) observation that visuo-spatial ability (rather than verbal) is the ability which most differentiates races. This also differs, as indicated by many studies (including that of Hansen-Strain & Strain), between the sexes, with males typically scoring higher than females (this being a well-documented general finding, e.g. see Bernard, Boyle, & Jackling, 1990).

Clearly, the above data would be shown to be of particular relevance to SLA research if, as discussed earlier, the proposed relationship between low spatial ability and superior language learning ability were to be conclusively demonstrated within the compensatory neuropsychological model proposed by Schneiderman &

Desmarais (1988). Indications of such sizeable differences between nations/races in visuo-spatial ability would then become of enormous L2 research interest, particularly in Asia. It would, for example, be a finding of some importance if the discrepancy between the Japanese and the Filipinos as indicated in the above tables, was actually observed to be related to aspects of language learning, thus strengthening the assumption of a neuropsychological substrate for language learning ability. Although at this time the evidence is not yet in, the availability of such rich cross-cultural data makes its collection all the more interesting and all the more essential, if we are to understand the influence of psychological variables in language learning.

A further example of the relevance of research-based information of this type to language learning becomes apparent when it is related to current teaching practices. In Japan, for example, although observational classroom data are difficult to obtain the overwhelming concensus from comparative education studies (e.g. Duke, 1986) and questionnaire investigations (e.g. see Aiga, 1990), is of extremely traditional language teaching methods being employed in high schools. As Aiga (1990, p. 140) observes "most of the average lesson period is spent on mechanical drills and on the teacher's explanations, rather than on communicative activities." Explanations for the continuance of grammar-translation and pronunciation drills as primary activities in Japanese schools usually centre on teachers teaching as they were taught, or the demands of university entrance exams. There is, however, also intuitive support for the view that the Japanese may be sufficiently different from Westerners so as to justify their adaptations of methods. Ito (1978, p. 214), for example, states (using the Germans merely as an example of Westerners), "The method suited to the German is not necessarily suited to the Japanese." Data to support such opinions are not cited, but if personality and mental ability findings on the Japanese are considered (and related to the demands of the schools to teach English to 94% of the school age population over the age of 16) then justification for employing tradition methodology can be derived.

Firstly, highly anxious introverted individuals are likely to be more comfortable with methods which do not force them into public performance, which may expose poor fluency and error production. And, as Lazarus, Tomita, Option Jr., & Kodama (1966) found when their experiment on cross-cultural anxiety was ruined by Japanese subjects being as anxious while watching a film on rice farming as they were when watching one on genital mutilation, the Japanese are unusually sensitive even to totally disinterested observation.

Secondly, evidence (e.g. Jenson, 1973, p. 6; 1974) indicating that low ability subjects more approximate high ability subjects on "associative learning ability" (rote learning) than they do on "conceptual learning ability" (abstract reasoning), suggests the former as more suitable for groups containing high proportions of low ability students who have high expectations. Experimental findings (e.g. Tinkham, 1990) also confirm that Japanese high school students are not only very good at rote learning, but they also have a more positive view of it than do American students. Japanese educationalists (e.g. Sato, 1978, p. 306) also appear to be more favourably disposed to rote learning than do most Westerns.

Personality variables may therefore partly account for present language teaching methodology in Japanese schools, and, in addition, help to explain why immediate post war attempts to introduce progressive teaching methods, with children taking an active role in the learning process, were seen as "unnatural" (Duke, p. 162) and were, consequently, short-lived. Efforts to introduce more communicative language teaching into Japan may consequently need to be redoubled if they are to surmount the hurdle presented by the Japanese personality and metal ability profile; only demonstrations of unquestionably superiority of communicative methods are likely to bring this about.

Solely in terms of personality, Singaporeans and subjects from Hong Kong appear to be much more similar to, for example, the British, than do the Japanese. Lynn (1977) does, however, indicate that Chinese Singaporeans exhibit high visuo-spatial ability comparable to that of the Japanese. Also, in terms of learning strategies, the findings of an L2 study aimed at improving vocabulary through imagery and grouping strategies (O'Malley, Chamot, Stewner-Manzanares, Russo, & Kupper, 1985) is perhaps instructive. Asian students' performance is summarised as follows: "Asian students [mainly from Southeast Asia] in the control group applied rote memorization strategies to the vocabulary task so successfully they outperformed the experimental groups who had been trained in what we perceived as more sophisticated strategies" (reported in O'Malley & Chamot, 1990, p. 165). With reference to rote memorization ability, Southeast Asians appear, therefore, to be not unlike the Japanese, and it raises the possibility that here too resistance to more communicative methods might persevere for this reason.

Clearly, at present the only substantive findings in the area are descriptive ones, the relationship of these variables to L2 learning requiring a good deal of further investigation. However, if as proposed earlier, personality variables are

demonstrated to play a significant role in explaining individual differences in aspects of language learning, then it would be illogical not to expect differences between nationalities (especially on the scale of the I-E variation between China and India, or on N between males of Japan and the U.K.), to be reflected in aspects of language classroom performance. Equally, if further support is forthcoming for the visuo-spatial-deficit/language-learning-ability hypothesis, the difference, for example, between, performance on visuo-spatial measures by the Japanese and Filipinos, might be seen to have explanatory value in a cross-cultural model of L2 Learning.

As is obvious, the many interesting relationships suggested by these data cannot now be considered as more than speculative; the area has been too long neglected for more to be possible at this stage. Yet it is also clear that cross-cultural studies of personality conducted within Eysenkian theory readily yield hypotheses relevant to SLA. In fact, insofar as variables such as I-E and verbal/visuo-spatial ability are demonstrated to be relevant to L2 learning, findings such as those described above suggest that a model of L2 learning which does not take cross-cultural differences into account will be incomplete.

The Personality of Teachers

It would perhaps not be appropriate to conclude a paper on personality and classroom learning without at least mentioning a little of the research on the personality of the teacher.

Early research (e.g. Evans & Wrenn, 1942) which suggested that extraverts were more successful teachers than introverts, was later questioned in studies on "educational seduction" or the "Dr Fox effect" (Naftulin, Ware, & Donnelly, 1973). In a meta-analysis of the literature on educational seduction, Abrami, Leventhal, and Raymond (1982) concluded that personal style, being entertaining and charismatic, etc., can merely seduce students into believing they have learned. They found "that instructor expressiveness had a substantial impact on student ratings but a small impact on student achievement. In contrast, lecture content had substantial impact on student achievement but a small impact on student ratings" (p. 446).

The current position (Murray, Rushton, & Paunonen, 1990), however, is that teacher personality traits are translated into specific classroom behaviours which are

validly reflected in student ratings. Although extraversion is indicated as a distinguishing characteristic of "good" teachers, another major finding from this research is that the specific personality traits which contribute to effective teaching, vary between course types, e.g. "neurotic extraverts" excel in large, lower-level, lecture classes, while neuroticism appears less desirable in graduate seminars.

It is notable that all of the above research has, however, been carried out with content teachers. Relating the personality of L2 teachers to language teaching variables is a little-tapped area of considerable research potential.

Conclusion

Evidence of the sensitivity of teachers to individual differences of students comes from the literature on teacher expectations and self-fulfilling prophecies. Here the conclusion (e.g. see Jussim, 1989) is that teachers' expectations predicted student performance "primarily because they were accurate" (p. 477). Results of a study of student-teacher interactions by Hummel-Rossi and Merrifield (1977), for example, also show that teachers are reasonably aware of the individual needs of learners and, where possible, respond to them. The general conclusion, however, is that personality dispositions are perceived less accurately than ability dimensions (e.g. see Jussim, 1989, p. 478).

Eysenck sees this problem in terms of the data available to teachers. Commenting on the relevance of personality variables to classroom practice, he points out that these are "interactions to which the best teachers have of course always been sensitive, but with which they have hitherto been able to deal only on an intuitive rather than a rational, experimentally supported basis" (1978, p. 153). Also, effects of intervention are not always obvious or in accord with common wisdom, e.g. individualized instruction (in comparison with the conventional lecture/examination method) intended to alleviate test anxiety has been observed to cause more anxiety than the original malady (Watson, 1988).

It is therefore necessary for the L2 researcher to explore such interactions and transmit whatever knowledge is available in the area of classroom practitioners. Both the theory and the available data make it appear desirable to conduct rigorous theory-based research on personality in an L2 context at a number of levels. The well-documented accounts of differences between races and nationalities suggests hypothesis formulation at this level is likely to be instructive, and conducting actual

case studies also has obvious relevance when individual differences are under discussion. Experimentation at levels between these extremes has also been described and positive findings are now forthcoming.

An additional, and extremely positive aspect of on-going research is that, in testing hypotheses from a developed theoretical model, findings from this research cannot only be used to inform the field of applied linguistics/EFL, but they can also feed back into personality theory. Clearly, a symbiotic relationship was always desirable, and, given the present reassessment of personality variables in the context of L2 teaching/learning, it is becoming a reality.

This development represents a radical change in the status of such variables in an L2 context, as until recently, there were few indications that the study of personality within the discipline of SLA merited serious consideration. However, now that the area is being seriously considered, findings of consequence are beginning to emerge. The study by Berry (forthcoming), which appears to be the first L2 investigation in the domain of personality to test hypotheses based on established personality theory and the first to use and appropriately validated personality scale derived from cross-cultural analysis, can be regarded as something of a breakthrough in the area. It certainly suggests exciting research-based prospects for the future, and the possibility of L2 personality research being truly "applied". Theory and research in the area are, in fact, now poised to make a contribution to practice.

It is, in conclusion, of course acknowledged that personality is only one of the many variables that play a part in explaining individual differences in SLA; it is, however, hoped that on-going research findings and restatements of theoretical positions will convince those who have been prematurely dismissive of the area, that it is worthy of sustained scientific attention. Unless "we assume that teaching can proceed just as well in the absence of any knowledge about the learning process, or individual differences relating to it..." (Eysenck, 1978, p. 169), then clearly it is.

The Author

Roger Griffiths has a Ph.D. in applied linguistics from Southampton University and is Associate Professor of English at Nagoya University of Commerce and Business Administration, Japan. He is also a Chartered Psychologist and Associate Fellow of The British Psychological Society.

REFERENCES

*ABRAMI, P.C., Leventhan, L., & Raymond, P.P. (1982). **Educational Seduction.** Review of Educational Research, 52, 446-464.*

*AIGA, Y. (1990). **Is Japanese English Education Changing?** Cross Currents, 2, 139-145.*

*AMARIA, R.P. & Leith, G.O.M. (1969). **Individual Versus Co-operative Learning.** Educational Research, 11, 193-199.*

*BARRETT, P. & Eysenck, S.G.B. (1984). **The Assessment of Personality Factors across 25 countries.** Personality and Individual Differences, 5, 615-632.*

*BERNARD, M.E., Boyle, G.J., & Jackling, B.F. (1990). **Sex-role Identity and Mental Ability.** Personality and Individual Differences, 11, 213-217.*

*BERRY, V. (forthcoming). **Personality Characteristics as a Source of Language Test Bias.** Paper presented at the Language Testing in Europe Seminar, August 1991, University of Jyvaskyla, Finland.*

*BLEASE, D. (1986). **Teachers' Personal Constructs and Their Pupils' Self-images.** Educational Studies, 12, 255-264.*

*BLEY-VROMAN, R. (1989). **What is the Logical Problem of Foreign Language Learning?** In S.M. Gass & J. Schachter (eds.), Linguistic Perspectives on Second Language Acquisition (pp. 41-68). Cambridge: CUP.*

*BROWN, H.D. (1973). **Affective Variables in Second Language Acquisition.** Language Learning, 23, 231-244.*

*BROWN, H.D. (1987). **Principles of Language Learning and Teaching. (2nd Edition).** Englewoood Cliffs, NJ: Prentice Hall Regents.*

*BURGER, J.M. (1986). **Personality.** Belmont, CA: Wadsworth.*

*BUSCH, D. (1982). **Introversion-extraversion and the EFL Proficiency of Japanese Students.** Language Learning, 32, 109-132.*

CAMPBELL, J.B. & Hawley, C.W. (1982). Study Habits and Eysenck's Theory of Extraversion-introversion. Journal of Research in Personality, 16, 139-146.

CATTELL, R.B. (1957). Personality and Motivational Structure and Measurement. Yonkers: New World.

CRONBACH, L.J. (1984). Essentials of Psychological Testing. Fourth Edition. New York: Harper & Row.

CRONBACH, L.J. (1990). Essentials of Psychological Testing. Fifth Edition. New York: Harper & Row.

de JONG, J.H.A.L. & Stevenson, D.K. (eds.), (1990). Individualizing the Assessment of Language Abilities. Clevedon: Multilingual Matters.

DIGMAN, J.M. (1989). Five Robust Trait Dimensions: Development, Stability and Utility. Journal of Personality, 57, 195-214.

DORNIC, S. & Ekehammer, B. (1990). Extraversion, Neuroticism and Noise Sensitivity. Personality and Individual Differences, 9, 989-992.

DUKE, B. (1986). The Japanese School: Lessons for Industrial America. New York: Praeger.

ELLIS, R. (1985). Understanding Second Language Acquisition. Oxford: OUP.

EYSENCK, H.J. (1967). The Biological Basis of Personality. Springfield, IL: Charles C. Thomas.

EYSENCK, H.J. (1970). The Structure of Human Personality. (3rd Edition). London: Methuen.

EYSENCK, H.J. (1976). The Measurement of Personality. Lancaster: Medical and Technical Publishers.

EYSENCK, H.J. (1978). The Development of Personality and its Relation to Learning. In S. Murray-Smith (ed.), Melbourne Studies in Education (pp. 134-181). Melbourne: Melbourne University Press.

EYSENCK, H.J. (1984). The Biology of Individual Differences. In V. Sarris & A. Parducci (eds.), Perspectives in Psychological Experimentation: Towards the Year 2000 (pp. 179-195). Hillsdale, NJ: Erlbaum.

*EYSENCK, H.J., & Eysenck M.W. (1985). **Personality and Individual Differences: A Natural Science Approach. New York: Plenum.***

*EYSENCK, H.J., & Eysenck, S.B.G. (1975). **Manual of the Eysenck Personality Questionnaire. London: Hodder and Stoughton.***

*EYSENCK, M.W. (1974). **Extraversion, Arousal, and Retrieval from Semantic Memory. Journal of Personality, 42, 319-331.***

*EYSENCK, M.W. (1977). **Human Memory: Theory, Research and Individual Differences. Oxford: Pergamon.***

*EYSENCK, M.W. (1970). **Anxiety, Learning, and Memory. Journal of Research in Personality, 13, 363-385.***

*EYSENCK, M.W. (1984). **A Handbook of Cognitive Psychology. London: Erlbaum.***

*EYSENCK, S.B.G. & Chan, J. (1982). **A Comparative Study of Personality in Adults and Children: Hong Kong vs England. Personality and Individual Differences, 3, 153-160.***

*EVANS, C. & Wrenn, C.G. (1942). **Introversion-extraversion as a Factor in Teacher Training. Educational and Psychological Measurement, 2, 47-58.***

*FLODERUS-MYRHED, B., Pederson, N., & Rasmuson, I. (1980). **Assessment of Heritability for Personality, based on short-form of the Eysenck Personality Inventory: A Study of 12,898 twins. Behavoural Genetics, 10, 153-162.***

*GENESSEE, F. & Hamayan, E. (1980). **Individual Differences in Second Language Learning. Applied Psycholinguistics, 1, 95-110.***

*GRIFFITHS, R.T. (1991). **Language Learning and the Japanese Personality: Hypotheses from the Psychological Literature. IUJ Conference Proceedings on Second Language Research in Japan. 20 October, Tokyo, Niigata: IUJ.***

*GRIFFITHS, R.T. & Berry, V. (in preparation). **The Japanese Personality: EPQ data on a student sample.***

*GRIFFITHS, R.T. & Sheen, R. (in preparation, a). **Japanese Visuo-spatial Abilities and Personality: Further evidence.***

*GRIFFITHS, R.T. & Sheen, R. (in preparation, b). **Embedded Figures in the Landscape: A Reappraisal of L2 Research on Field Dependence/Independence.***

*HANSEN-STRAIN, L. & Strain, J. (....) **Learning Styles of the Japanese Second Language Learner: Field Sensitivity, Cognitive Tempo, Test Anxiety, Modality Preference** ..... 257-275.*

*HOWARTH, E., & Eysenck, H.J. (1968). **Extraversion, Arousal, and Paired-associate Recall.** Journal of Experimental Research in Personality, 3, 114-116.*

*HUMMEL-ROSSI, B., & Merrifield, P. (1977). **Student Personality Factors related to Teacher Reports of their Interactions with Students.** Journal of Educational Psychology, 69, 375-380.*

*ITO, K. (1978). **Traditional Methods and New Methods: A study on the Methods suited for the Japanese.** In I. Koike, M. Matsuyama, Y. Igarashi, & K. Suzuki (eds.), The Teaching of English in Japan (pp. 204-219). Tokyo: Eichosha.*

*IWAWAKI, S., Eysenck, S.B.G., & Eysenck, H.J. (1980). **Japanese and English Personality Structure: A Cross-cultural Study.** Psychologia, 23, 195-205.*

*JENSON, A.R. (1973). **Educability and Group Differences.** London: Methuen.*

*JENSON, A.R. (1974). **Interaction of Level I and Level II Abilities with Race and Socioeconomic Status.** Journal of Educational Psychology, 66, 99-111.*

*JENSON, A.R. & Reynolds, C.R. (1982). **Race, Social Class and Ability Patterns on the WISC-R.** Personality and Individual Differences, 3, 423-438.*

*JUSSIM, L. (1989). **Teacher Expectations: Self-fulfilling Prophecies, Perceptual Biases, and Accuracy.** Journal of Personality and Social Psychology, 57, 469-480.*

*LAZARUS, R.S., Tomita, M., Opton Jr., E, & Kodama, M. (1966). **A Cross-cultural Study of Stress-reaction patterns in Japan.** Journal of Personality and Social Psychology, 4, 622-633.*

*LEITH, G.O. (1969). **Learning and Personality.** In W.R. Dunn & C. Holroyd (eds.), Aspects of Educational Technology, Vol. 2 (pp. 101-110). London: Methuen.*

*LEITH, G.O. (1972). **The Relationship between Intelligence, Personality and Creativity under Two Conditions of Stress.** British Journal of Educational Psychology, 42, 240-247.*

*LEITH, G.O. (1973). **The Effects of Extraversion and Methods of Programmed Instruction on Achievement.** Educational Research, 5, 150-153.*

*LEITH, G.O. (1974). **Individual Differences in Learning: Interactions of Personality and Teaching Methods.** Personality and Academic Progress Proceedings, 14-25. London: Association of Educational Psychologists.*

*LEITH, G.O & Trown, E.A. (1970). **The Influence of Personality and Task Conditions on Learning and Transfer.** Programmed Learning, 7, 181-188.*

*LEITH, G.O., & Wisdom, B. (1970). **An Investigation of the Effects of Error Making and Personality on Learning.** Programmed Learning, 7, 120-126.*

*LONG, M.H. (1990). **Maturational Constraints on Language Development.** Studies in Second Language Acquisition, 12, 251-285.*

*LONG, M.H. & Porter, P.A. (1985). **Group Work, Interlanguage Talk, and Second Language Acquisition.** TESOL Quarterly, 19, 207-228.*

*LYNN, R. (1977). **The Intelligence of the Chinese and Malays in Singapore.** Mankind Quarterly, 18, 125-128.*

*LYNN, R. (1987). **The Intelligence of the Mongoloids: A Psychometric, Evolutionary and Neurological Theory.** Personality and Individual Differences, 8, 813-844.*

*McCULLERS, J.C., Fabes, R.A., & Moran III, J.D. (1987). **Does Intrinsic Motivation Theory Explain the Adverse effect of Rewards on Immediate Task Performance?** Journal of Personality and Social Psychology, 52, 1027-1033.*

*McDONOUGH, S.H. (1986). **Psychology in Foreign Language Teaching. (2nd Edition).** London: Unwin Hyman.*

*MURRAY, H.G., Rushton, J.P., & Paunonen, S.V. (1990). **Teacher Personality Traits and Student Instructional Ratings in Six Types of University Course.** Journal of Educational Psychology, 82, 250-261.*

*NAFTULIN, D.H., Ware, J.E. Jr., & Donnelly, F.A. (1973). **The Doctor Fox Lecture: A Paradigm of Educational Seduction.** Journal of Medical Education, 48, 630-635.*

*NAGPAL, M., & Gupta, B.S. (1970). **Personality, Reinforcement and Verbal Operant Conditioning.** British Journal of Psychology, 70, 471-476.*

*NAIMAN, N., Frolich, M., & Stern, H.H. (1975). **The Good Language Learner.** Toronto: The Ontario Institute for Studies in Education.*

*NAIMAN, N., Frolich, M., Stern, H.H., & Todesco, A. (1978). **The Good Language Learner.** Research in Education Series No. 7. Toronto: The Ontario Institute for Studies in Education.*

*NOVOA, L., Obler, L., & Fein, D.A. (1988). **A Neuropsychological Approach to Talented Second Language Acquisition - A Case Study.** In L.K. Obler & D.A. Fein (eds.), The Exceptional Brain: Neuropsychology of Talent and Special Abilities, (pp. 294-302). Cambridge: CUP.*

*O'MALLEY, J.M., Chamot, A.U., Stewner-Manzanares, G., Russo, R., & Kupper, I. (1985). **Learner Strategy Applications with Students of English as a Second Language.** TESOL Quarterly, 19, 285-296.*

*O'MALLEY, J.M., & Chamot, A.U. (1990). **Learner Strategies in Second Language Acquisition.** Cambridge: CUP.*

*REVELLE, W., Humphreys, M.S., Simon, L., & Gilliland, K. (1980). **The Interactive Effects of Personality, Time of Day and Caffeine:** A Test of the Arousal Model. Journal of Experimental Psychology: General, 109, 1-31.*

*ROSE, R.J., Koshenvuo, M., Kaprio, J., sarna, S., & Langinvainio, H. (1980). **Shared Genes, Shared Experiences, and Similarity of Personality;** Data from 14,288 Adult Finnish Co-twins. Journal of Personality and Social Psychology, 54, 161-171.*

*ROYCE, J.R. (1988). **The Factor Model as a Theoretical Basis for Individual Differences.** In S.H. Irvine & J.W. Berry (Eds.), Human Abilities in Cultural Context, (pp. 147-165). Cambridge: CUP.*

*RUBIN, J. (1975). **What the "Good Language Learner" can teach us.** TESOL Quarterly, 9, 41-51.*

*SATO, H. (1978). **Production:** On Classroom Activities. In I. Koike, M. Matsuyama, Y. Igarashi, & K. Suzuki (Eds.), The Teaching of English in Japan. (pp. 298-308). Tokyo: Eichosha.*

SCHNEIDERMAN, E. & Desmarais, C. (1988). **A Neuropsychological Substrate for Talent in Second Language Acquisition.** In L.K. Obler & D. Fein (eds.), The Exceptional Brain: Neuropsychology of Talent and Special Abilities . (pp. 103-126). New York: Guilford.

SKEHAN, P. (1989). **Individual Differences in Second-language Learning.** London: Edward Arnold.

SMITH, I. MacFARLANE (1964). **Special Ability.** London: University of London Press.

STERNBERG, G., Risberg, J., Warkentin, S., & Rosen, I. (1990). **Regional Patterns of Cortical Blood Flow Distinguish Extraverts from Introverts.** Personality and Individual Differences. 7, 663-673.

SWAIN, M. & Burnaby, B. (1970). **Personality Characteristics and Second Language Learning in Young Children:** A Pilot Study. Working Papers in Bilingualism, 11, 76-90.

TINKHAM, T. (1990). **Rote Learning, Attitudes, and Abilities:** A Comparison of Japanese and American Students. TESOL Quarterly, 24, 695-698.

WANKOWSKI, J. (1978). **Personality Dimensions of Students and Some Educational Implications of Eysenck's Theory of Extraversion and Neuroticism.** University of Birmingham Educational Survey Research Report. Birmingham: University of Birmingham.

WATSON, J.M. (1988). **Achievement Anxiety Test:** Dimensionality and Utility. Journal of Educational Psychology, 80, 585-591.

WERNER, E.E. (1979). **Subcultural Differences in Ability, Achievement, and Personality Factors among Oriental and Polynesian Children on the Island of Kauai.** In L. Eckensberger, W. Lonner, & Y.H. Poortinga (eds.), Cross-cultural Contributions to Psychology. Lisse: Swets & Zeitlinger.

WESCHE, M., Edwards, H., & Wells, W. (1982). **Foreign Language Aptitude and Intelligence.** Applied Psycholinguistics, 3, 127-140.

WITKIN, H.A., Oltman, P.K., Raskin, E., & Karp, S.A. (1971). **A Manual for the Embedded Figures Test.** Palo Alto, Ca.: Consulting Psychologists Press.

THE CONTRIBUTION OF SLA THEORIES AND RESEARCH TO TEACHING LANGUAGE[1]

Andrew D. Cohen, Diane Larsen-Freeman and Elaine Tarone

Abstracts

This paper highlights some of the areas in which SLA theories and research have contributed to language teaching. The paper notes that while results of SLA research may have contributed to our understanding of language learning, insights from such research may have little direct effect on classroom instruction. One explanation for this lack of effect is that the SLA research agenda is not necessarily that of second-language-teaching (SLT) research. The paper culls from the SLA research literature six areas in which SLA (and SLT) research findings have had or could have impact on teachers' awareness: comprehensible input, focus on form, correction of speaking errors, pronunciation, speech act sets, learning strategies and factors influencing language learners. It is concluded that a knowledge of SLA research findings helps to inform teachers' decisions, even if these findings are not directly applicable to the classroom, while some of the concepts and tools developed in the process of research on SLA may be directly useful to teachers in conducting needs assessment.

Introduction

The purpose of this paper is to highlight some of the areas in which SLA theories and research have contributed to language teaching. The task is not straightforward in that while the results of SLA research may have contributed to our understanding of language learning, insights from such research may have little direct effect on classroom instruction (Larsen-Freeman 1990). One explanation for this lack of effect is that the SLA research agenda is not necessarily that of second-language-teaching (SLT) research. In SLT research, the interest is directed towards how and why classroom interactions or features contribute to learning opportunities.. SLA research has tended to focus more (though not exclusively) on untutored language learning and on out-of-class contexts, and what is minimally necessary for SLA to take place.

What is minimally necessary for SLA to occur in untutored contexts is not necessarily what we should emulate in the classroom. For example, some SLA researchers have pointed out that SLA occurs without a focus on form, often through a process of communicative interaction. While such an observation may be valid, it does not follow that practitioners should eliminate such a focus in their classrooms. One would hope that effective teaching would accelerate the natural process, or else why would one seek a teacher in the first place? Accelerating SLA might very well involve helpful intervention in the "natural" process just as a medical doctor intervenes so as to speed up the natural healing process of patients.[2]

There are two broad ways in which SLA research has contributed to language teaching: (1) the findings of such research have enhanced teachers' general understanding of second-language acquisition, and (2) many of the concepts and tools developed in researching SLA have proven useful to teachers engaged in the process of needs assessment.

Even though observations of natural learning may not always translate directly into classroom practice, there is no question that findings from SLA research can do much to enhance teachers' understanding of second-language acquisition. With enhanced understanding, teachers can make more informed decisions and build upon, rather than work against, learners' natural inclinations (Larsen-Freeman 1983). From this conviction, then, we have culled from the SLA research literature, six areas in which SLA (and SLT) research findings have had or could have impact on teachers' awareness: comprehensible input, focus on form, correction of speaking errors, pronunciation, speech act sets, learning strategies and factors influencing language learners. It is these six areas which constitute the primary focus of this paper.

1. Research Findings

Comprehensible Input

A hypothesis that has had an impact on language teaching, despite the controversy surrounding it, is that learners move most rapidly toward mastery of a language by acquiring it through comprehensible input (Krashen 1985, 1989). This hypothesis states that more comprehensible input results in more language acquisition, that language teaching methods containing more comprehensible input are more effective, and that language development occurs more effectively without

formal instruction focusing on conscious learning. In essence, both children and adults are seen to be able to "acquire" a second language (Lightbown 1985). The message to language teachers has been that rather than attempting to teach the numerous structures of the language, they should focus more on making the language available to the learners for them to acquire forms that are salient to them at their current level of language development.

Recent reviews of the SLA literature would question the minimizing of the effects of instruction on language learning. A review by Long (1988), for example, found that formal instruction does have positive effects on SLA processes. Formal target-language instruction has been found to speed up the rate at which learners acquire the language forms and also to result in a higher ultimate level of attainment (Ellis 1989). In fact, Long (1988) argues that it may be impossible to attain full native speaker competence without instruction. Precisely how and why instruction is facilitative of SLA is an issue we feel is appropriate for an SLT research agenda.

It has also been claimed that input alone is not sufficient for learning to take place, but rather that the learner needs opportunities for negotiation, which has been seen to lead to appropriate modifications in input complexity and amount of redundancy (see, for example, Long 1983). Recent studies have added qualifiers to the issue of negotiation. Ehrlich, Avery, & Yorio (1989), for example, found that meaning negotiations did not automatically benefit Japanese ESL learners interacting in pairs with native speakers, working on a problem-solving picture-drawing task. Rather, the success or failure of meaning negotiations in providing comprehensible input was seen to depend in part on the discourse strategy employed by the interlocutor: **skeletonizing** (i.e., providing only the bare events of a narrative) provided greater opportunity for comprehension than did **embroidering** (i.e., an expanded, embellished description).

Skeletonizers were more likely to repeat previously mentioned information, perhaps breaking it up into smaller units, while embroiderers were more likely to provide new information, in some cases with an inordinate amount of detail. If natives provided skeletonized discourse, the nonnatives were better able to locate the sources of their non-understanding than in embroidered discourse, where each additional detail rendered it more deeply embedded and therefore less understandable. These researchers concluded that if the interaction was to stimulate change in the learners' interlanguage system, then the learners needed to recognize the precise nature of their non-understandings.

This finding runs counter to an assumption in the literature that the mere quantity of meaning negotiations within a discourse is an accurate predictor of the quantity of comprehensible input that results. Hence, such negotiations need to be analyzed within a discourse framework to explain their role in creating comprehensible input. Indeed, many complex factors may determine the usefulness of interactive tasks in providing comprehensible input to learners. Yule and Macdonald (1990) found that in tasks presenting specific referential conflicts which needed to be resolved, higher proficiency learners who were assigned a dominant role often refused to engage in interactive cooperation with lower proficiency learners, sometimes even changing the task rather than negotiating. In such cases, the amount of negotiation and presumably of comprehensible input to the learner improved substantially when higher proficiency learners were placed in nondominant roles. The research by Ehrlich et al. and Yule and Macdonald suggests why particular communicative tasks may only work for certain kinds of discourse situations and not for others, and that the whole area is much more complicated than previously thought.

Focus on Form

Closely aligned with the discussion of comprehensible input and meaning negotiation is the issue of focus on form. More traditional language teaching methods have often put a premium on drilling of grammatical forms as a way of teaching them to learners. Yet the effectiveness of grammar instruction depends on the sequencing of grammar rules and the careful assessment of learner readiness (Pica 1989). At least in the case of grammatical morphemes and other forms that have been studied, it has been found that instruction does not change the natural sequence of acquisition, although it can accelerate the movement across stages of development. Furthermore, "practice does not make perfect" in that even though there are acquisition sequences, acquisition is not simply linear or cumulative, and having practiced a particular form or pattern does not mean that the form or pattern is permanently established. Learners appear to forget forms and structures which they had seemed previously to master and which they had extensively practiced. One explanation for this is that the encountering of new forms causes a restructuring of the learner's whole language system, which may result in simplification (however temporary) in some other part of the system (Lightbown 1985:177).

With regard to drills, it has been seen that they seem to have generally mixed success in getting learners to internalize the correct grammatical forms. Teachers have often noticed that even if learners are able to demonstrate reasonable control of given structures in practice, they fail once they are called upon to use the same structures in communication. Research documenting such task-related shifts in accuracy of grammar as well as of pronunciation is summarized in Tarone (1988). One theory is that accuracy shifts as learner attention shifts from form to meaning. There is some research suggesting that especially in early stages, learners have problems attending to form and meaning at the same time. In a study of 202 English-speaking students of Spanish, the participants were given four tasks--attending to meaning alone, attending simultaneously to meaning and to an important (i.e., communicative) lexical item, attending to meaning and to a grammatical functor, and attending to meaning and a verb form (VanPatten 1990). Recall of meaning was lowest on the last two tasks, where learners focused simultaneously on form and meaning. Yet classroom teachers often request that students focus not only on the content of the message but on its grammatical form as well.

Given this problem, researchers have begun to take a compromise position supported by SLA theory which is that learners should be led to notice grammatical features in the input, compare what they have noticed with what they produce in their current interlanguage, and then eventually integrate the new features into their interlanguage when they are ready (Ellis 1990). This is a departure from the approach that would have them producing correct grammatical forms from the start of the course. Another approach consistent with SLA findings is one that recognizes that all language units have three dimensions (form, meaning, and use) and that it is the teacher's task to systematically focus upon only one of these dimensions at a time, shifting the focus as the needs of the learners change (Larsen-Freeman, 1991).

Correction of Speaking Errors

While SLA researchers have not done extensive research on the effects of error correction on speaking (far more work has been done on writing), the evidence that is available would suggest that explicit error correction may be ineffective in changing language behavior (Lightbown 1985). In an extensive case study on his own learning of Portuguese, for example, Schmidt found corrections by others not to be too helpful because he was not always aware that he was being corrected nor

did he necessarily understand the problem (Schmidt & Frota 1986). The researchers concluded that learners need to notice gaps consciously in order to make progress in the target language.

It has been posited that we progress in a target language by testing hypotheses about how the language works, on the basis of inferences based on previous knowledge (Schachter 1983), and that learners depend on negative input to verify hypotheses about whether their utterances are comprehensible, grammatically correct, or situationally appropriate (Schachter 1984). Such input includes not only explicit correction, but also confirmation checks (i.e., confirmation elicited by the speaker that the utterance was correctly understood or heard) and requests for clarification (new information or a rephrasing of what was already said). Some argue that even a learner's failure to understand an utterance can provide negative input.

The problem is that whatever approach is selected by the teacher, Allwright (1975), Krashen (1982), and others would argue that the correction of oral errors will probably have limited or no effect if learners: (1) are not focused on the form of their message (i.e., its vocabulary, grammar, or pronunciation), (2) do not have enough time to consider the correction, (3) do not have adequate knowledge of the area being corrected to benefit from the correction, (4) have too little knowledge about how the language works to know what question to ask to get clarification or do not have adequate proficiency to understand the teacher's explanation of what they did wrong. Nevertheless, the research basis for these pronouncements is limited. For example, a major survey of studies on oral feedback (Chaudron 1988:136-153) found a lack of research concerning the impact of corrections on learners--i.e., what learners actually do with the corrections, if anything.

Two Hebrew University student seminar projects would suggest that correction of oral target-language utterances may not even be attended to at all--or only ineffectively, and that even repeated and blatant corrections may not "take." Alamari (1982), for example, looked at the way in which 26 advanced adult Hebrew second-language learners in four classrooms related to their teacher's oral correction. She recorded each instance in which a learner was corrected, and then approached the learners at the break in order to ask them what they did when their oral language was corrected in class. Although all the learners said that they wanted to be corrected and almost all said they took teacher corrections seriously, about 20% reported not paying attention to the corrections and only 15% said that they wrote down the correction in their notebooks. Mostly, they reported repeating

the correction to themselves. Given our largely intuitive sense that on-the-spot repetition alone may be of limited value, we might question whether it benefited those learners who reported using it as a technique.

Rosenstein (1982) conducted an interventionist study as teacher of a 100-hour university EFL course in spoken English. He collected two-minute segments of spoken language from each student in each of six class sessions (12 minutes in all) as a pretest and then another twelve minutes of speech as a posttest. An analysis of the transcriptions of the pretest allowed the teacher/investigator to assign all learners an overt error as their "public error" in need of eradication. He also assigned each learner a covert or "secret" error, one that they did not know about. He made sure that this covert error was another student's overt error in order to see if learners would learn from overhearing another student corrected on "their" error. The learners were corrected repeatedly on their overt error. Of the eight students for whom he had complete data, two showed significant improvement in their public error and one in her secret error at the end of the semester. Another two students showed improvement in their public error and one student in her secret error, but these findings were not at a level of statistical significance. The others showed modest or no improvement in their public and secret error. Rosenstein credited the level of success attained to his general discussions with the learners as to why they made errors, individual discussions with them about their particular public error and explanations for it, written assignments regarding the error, and immediate correction of the public error when occurring in speech. Yet his success was still only about 50% for public errors and perhaps 20% for secret errors. Furthermore, for those two students who did experience improvement in their public error, the errors may have been ones that the learners were just about to acquire anyway in sequence.

The reason why at least half the students managed to emerge from the treatment with little or no improvement can perhaps best be found in the Alamari study: the learners simply were not paying attention to the corrections, not paying attention well enough, or paying attention but not making an effort to remember--as, for example, by efficiently recording the feedback that they received for future reference. Another possibility is that the treatment did have a delayed effect which went unmeasured in this study (cf. Yule, Powers, & Macdonald 1991).

Holley and King (1971) in a study of first, second, and third-year college students of German found that if their teachers gave them from five to ten seconds' wait time in order for them to check out what they had said and to self-correct

before teachers jumped in to correct, the learners were able to spot and rectify more than half of their errors. This study, which involved the use of videotapes, found that the teacher's pause and the non-verbal expectation of student performance created a class atmosphere conducive to student self-correction. The time interval did not produce tension and did not slow the tempo of the lesson noticeably. To the contrary, it was found that teacher correction, explanation, and restatement of the questions took up as much or more class time than extra seconds of silence.

Another study, conducted at the University of California at Los Angeles, focused just on learners' ability to spot the oral language errors that they made (Schlue 1977). Three intermediate college ESL students were audio-taped for 15 minutes once a week and then listened to their tapes for 45 minutes, over a ten-week period. It was found that the students were able to spot their own errors only 25-40% of the time. In other words, they were oblivious to at least 60% of their errors, with their attention to form decreasing the more they wanted to communicate. When they were able to spot their errors, they were able to correct most of these. Hence, one message to the learner and to the teacher would be to make sure that the learner is afforded an opportunity to perform on-line correction of errors while speaking. This means that the teacher gives the learner time to self-correct rather than providing immediate correction. Another pedagogical implication would be that activities should have a clear intent: accuracy or fluency.

More recent research would suggest that if learners are allowed to make errors within the framework of highly controlled transfer exercises, then correction may have a positive benefit. Thirty-two English-speaking college students of introductory French received one of two treatments dealing with eight English-French negative transfer errors (Tomasello & Herron 1989). In the "Garden Path" condition, students were given a sentence that was likely to be mistranslated because of the difference between the two languages. When the inevitable error occurred, the teacher translated for the students, thereby calling attention to the negative transfer problem that resulted in the error. In the control condition, the students were given the correct French form, told that it was different, and thus not given an opportunity to commit a transfer error. Student learning of the non-transferable form was assessed three times throughout the course of the semester, and each time performance was better in the Garden Path condition. The conclusion was that students learned best when they produced an erroneous hypothesis and received immediate feedback--i.e., got to compare their system with that of the native. This finding confirms the theoretical position by Schachter (mentioned above) regarding the need for explicit negative input.[3]

It also underscores a more general SLA finding that has contributed to increasing teacher tolerance for learner errors. Where it was once thought that errors should be prevented at all cost, it is now understood that error commission is part of the learning process. As such, teacher correction needs to be judicious. This could mean, for example, correcting when the learners are ready for the corrections and have adequate knowledge about the structures involved (as in the "Garden Path" study), or when they have time to digest the corrections.

Pronunciation

The SLA literature has perhaps had some impact on the way in which teachers relate to pronunciation accuracy in the classroom. Whereas instruction in the 1940s and 1950s was sometimes built around accurate mimicry of target-language sounds, especially in the heyday of the audiolingual method, researchers have come to find that the accuracy of pronunciation varies when learners are asked to perform different tasks (Dickerson 1975).

One SLA researcher went so far as to demonstrate how increased and exclusive attention to mimicry of foreign-language sounds without knowing what they mean can lead to more accurate pronunciation. Twenty English-speaking university students, ranging in age from 19 to 22, were taught exclusively phonetic material on Chinese, Japanese, and Eskimo, using an eighteen-hour videotaped program for individualized instruction (Neufeld 1977). As a final test, the learners had to produce ten statements which were rated by native speakers. In the case of Japanese, the three raters rated three of the nonnatives as unmistakably native and six as native with traces of linguistic interference from English. In the case of Chinese, one was rated native and seven natives with traces of interference.[4] The finding would suggest that learners far after puberty are capable of achieving native-like pronunciation in a language if that is all that they focus on and if they do not know what they are saying. Yet the reality of the classroom is that a lot more is going on than instruction in the phonetics of the language.

As in the case of grammar, students may actually exhibit control of sounds in practice situations in the classroom, just as they can exhibit what appears to be control of grammatical structures. Then in actual communicative situations when they are not focusing on the forms and not monitoring for correctness, their control seems to break down. This was Cohen's observation when he taught English pronunciation to foreign students at the University of California at Los Angeles.

Some of the students could sound quite nativelike in mimicry exercises but in free speech sounded clearly nonnative. Research on task-related shifts in accuracy of pronunciation (Tarone 1988) shows that this phenomenon is widespread and to be expected.

Speech Act Sets

Another area in which SLA research has had an impact, however indirect, on language teaching is that of speech act sets. It has become increasingly clear to researchers that learners of a language may lack even partial mastery of such speech act sets' and that this lack of mastery may cause difficulties or even breakdowns in communication. Early empirical research on speech act sets (e.g., Cohen & Olshtain 1981) was in part prompted by a realization that although transfer occurs at the sociocultural level, few if any contrastive studies were systematically characterizing such phenomena (Schmidt & Richards 1980, Riley 1981, Loveday 1982). SLA research has helped to provide empirical descriptions of speech acts such as requests, compliments, apologies, complaints, refusals, and expressions of gratitude (see Wolfson & Judd 1983, Wolfson 1989).

In recent years, teachers have been encouraged to give attention in their instruction to speech act sets that are likely to be called upon in given speech situations. Prior to the advent of SLA research on the topic, teaching materials dealing with speech acts had for the most part been constructed largely in the absence of empirical studies to draw upon. They had relied on the curriculum writer's intuition and could best be characterized as reflecting a high level of simplicity and generality. Popular English-foreign-language textbooks treated speech act sets such as "apology" rather simplistically. For example, emphasis was almost exclusively on phrases used for the expression of an apology: **sorry, I'm sorry, I'm very sorry,** etc. Brief reference was made to other apology strategies, but without underlying principles for when to use what. No effort was made to analyze the apology speech act set into its semantic formulas, i.e., the various verbal realizations of an apology (see, for example, Blundell, Higgens, & Middlemass 1982; Berry & Bailey 1983; Swan & Walter 1985).

Empirical studies concerning the nature of various speech acts in a variety of languages and cultures have been steadily accumulating over the last few years (with respect to the apology speech act set, for example, see Cohen, Olshtain, & Rosenstein 1986; Olshtain & Cohen 1989, 1990). As a result there is a growing

source of empirical data on the strategies for performing these acts. Hence, SLA research has provided an opportunity for teachers and textbook writers to move from general, intuitively-based materials to more specific, empirically-based ones, which take into account variation resulting from differing levels of formality, severity of the incident, setting and interlocutors, and numerous other variables.

Whereas it is still too early to draw definitive conclusions as to the efficacy of formally teaching speech acts, a study by Olshtain and Cohen (1990) would suggest that the fine points of speech act behavior such as (1) types of intensification and downgrading, (2) subtle differences between speech act strategy realizations, and (3) consideration of situational features, can be taught in the foreign-language classroom. Yet the results of their study left Olshtain and Cohen skeptical regarding the effects of such instruction on the proficiency level of speech act behavior. In a study of nine female Japanese ESL learners tutored in complimenting and responding to compliments and nine untutored students of similar background, Billmyer (1990) collected data of weekly meetings between matched pairs of natives and nonnatives. Participants in both groups were asked to perform compliment-inducing tasks such as showing photos of home and family, reporting accomplishment, visiting each other's homes, teaching each other a proverb, and displaying a new item of apparel. It was found that tutored learners produced a greater number of norm-appropriate compliments, produced spontaneous compliments (unlike the untutored group), used a more extensive repertoire of semantically positive adjectives, and were far more likely to deflect the compliment in their reply than were their untutored peers. Billmyer concluded that formal instruction concerning the social rules of language use given in the classroom can assist learners in communicating more appropriately with natives outside of the classroom.

Although the results of these two studies are encouraging, it is still the case that EFL learners need to acquire not only a new repertoire of realization patterns in the new language but also change some of their speech act behavior, which can take a long time. Thus, perhaps the best that teachers can hope to achieve in the classroom is to create among learners a level of residual awareness so that they will be less prone to commit pragmatic failure both as producers and receivers of speech act behavior, and come to approximate native behavior more rapidly (Olshtain & Cohen 1990:57).

Learner Strategies and Factors Influencing Language Learners

Of all the contributions of SLA research to language pedagogy, one of the greatest has been the significance it has ascribed to the learning process. From the initiation of SLA research, investigators have been interested in the strategies learners make use of to acquire an L2. Early on, it was recognized that learners invoked strategies such as inferencing, hypothesis formation and testing, and using formulaic speech (first as routines and later as more analyzed patterns). With the pioneering work of Rubin (1975), attention was directed to the particular learning strategies of good language learners. Since that time a good deal of research has been devoted to continuing to identify and classify learning strategies and to determining if they are teachable (see, for example, Cohen 1990). Certainly the potential impact on language teaching is tremendous if teachers subscribe to Wenden's (1985) contention that they should not be content to regard their subject matter simply as language, but rather should be engaged in helping learners learn for themselves.

It is not only in the area of learning strategy preference that SLA research has contributed to rethinking how instruction should be shaped. Whereas at one time research on learner characteristics primarily dealt with aptitude, attitudes and motivation, SLA research the past two decades has also investigated personality factors, cognitive styles, hemispheric specialization, memory, interests, prior experience, birth order, etc. (Larsen-Freeman & Long 1991). While no specific pedagogical techniques can be prescribed based on the evidence that has been adduced thus far, the very fact that there is such variety among learners should underscore the need for teachers to see students as individuals and to work in a way as to take into consideration the diversity of backgrounds in their classes. Brown (1991), in fact, provides practical suggestions for individual learners according to their intellectual strengths, cognitive styles, and motivational level.

2. Research Tools

This last finding from SLA research--that there is tremendous individual variation among second-language learners--leads directly to the second broad contribution which SLA research and theory have made to language teaching, as noted at the outset of this paper: that of providing concepts and tools used in the process of SLA research which can also be applied to the assessment of learners' needs. Second-language teachers must always innovate to some extent. No one

syllabus or set of materials ever fits a group of learners exactly or does any general SLA finding. As a result, second-language teachers themselves are always engaged in a process of research on second-language acquisition--admittedly, research at a very local level, but research nonetheless. That is, teachers themselves need to be able to identify what it is that particular students and groups of students know of the L2 in order to decide how to proceed next. In that process of needs assessment, L2 teachers can be aided by concepts and tools contributed by SLA theories and research.

Tarone and Yule (1989) set out a variety of these concepts and tools, and outline their uses in L2 classroom needs assessment in greater detail than we can provide here. However, some examples may illustrate the richness of this resource for the classroom teacher. Tasks used to elicit data in SLA research are often appropriate for the elicitation of language in the language classroom, both for the purposes of assessment and (often) for the purpose of instruction and practice. For example, in a communicative task all learners can be given the same content to convey to a listener who needs that content in such a way that the content and form are controlled, but the language is communicative. In such tasks teachers can easily compare one learner to another, and learners to native speakers, under the same conditions.

As another example, a task used in investigating learner confidence (Yule, Yanz, & Tsuda 1985) can be used in combination with traditional multiple-choice grammar tests to provide the teacher with a measure of learners' confidence in their answers on that test. After choosing an item on a multiple-choice test, learners are asked to indicate on a scale from 1 to 5 how confident they are of the correctness of their choice. Thus, the teacher can identify those learners who are confident and correct, those who are confident and wrong, and those who are correct but not sure. Such information is surely useful to teachers interested in determining when it is time to move on in the syllabus.[6] As a final example, research on sociolinguistic skills of second-language learners has made use of recorded interactions and interviews. Shirley Brice Heath (1986) has proposed a procedure for classroom instruction and assessment which involves helping students in second-language contexts to record their own interactions with friends, family, fellow workers and so on. These language data are then discussed in class. The students may bring up whatever they want to discuss, but usually the focus is on the language as used for social interaction. Only the learners know who the speakers were, what the social context was, and what the outcome of the interaction was. They then function as the experts, and the teacher functions as a consultant to help interpret the language brought to class.

Sometimes frameworks used in SLA research (including frameworks borrowed from other disciplines and adapted for research purposes, as well as frameworks developed solely for SLA research) can also be useful frameworks for thinking about classroom needs assessment and instruction. One of the most useful is Canale and Swain's (1980) concept of **communicative competence.** Tarone and Yule (1989) suggest that grammatical competence is the ability to produce **accurate** language, sociolinguistic competence is the ability to produce **appropriate** language, and strategic competence is the ability to **successfully transmit information** in a language. This construct has proved very useful to teachers interested in evaluating their students' L2 skills, since learners may be quite proficient along one dimension, and not others. Another useful framework might be based on Speech Accommodation Theory (e.g., Beebe & Giles 1984), a current model for the analysis of sociolinguistic interaction. This model suggests that when speakers desire the social approval of their listeners, their speech patterns tend to converge toward those of their listeners. Conversely, when speakers desire to emphasize their differing group membership, their speech patterns tend to diverge. Such a model might provide a useful framework for teachers interested in analyzing the reasons for progress (or lack of it) in the SLA of particular classroom learners.

Our point here is that, quite apart from firm findings which SLA research can offer classroom teachers--general results about stages of SLA, for example--there are also concepts and tools being developed in SLA research and theory-building which may be directly useful to teachers involved in the process of needs assessment and classroom instruction.

Conclusions

It should be apparent from the topics discussed in this paper that SLA research has made learners and learning central, and in some ways has thus contributed to a shift in focus from how teachers teach to how and what learners learn. Yet while the focus on learners and learning is important, it is not one that always offers straightforward answers to teachers. We have suggested certain pedagogical practices that have been a direct or indirect result of SLA research[7]. We feel that a knowledge of SLA research findings helps to inform teachers' decisions, even if these findings are not directly applicable to the classroom. With regard to comprehensible input, it has been noted that negotiation for meaning may be important but that such negotiation is a complex matter. Research would further suggest that while a focus on grammatical form is valuable, it is important to focus

on meaning and on appropriate language use as well. As concerns whether the correction of speaking errors "takes," it is likely that learners have to be ready for the corrections, have to have adequate knowledge about what is being corrected, and have to have ample time to digest the corrections. With regard to pronunciation, accuracy may vary by task, possibly with poorer pronunciation resulting in situations where the learner is focusing more on conveying meanings than the correct sounds.

The correct realization of speech act sets poses a real challenge for the learner, wherein an awareness of the variables involved may help lead the learner to more successful speech act comprehension and production. Learner strategies are seen as important in that they help learners to help themselves, thus freeing teachers to be in more of a support role. Also, SLA research has helped to establish the real need to take into account the level of diversity among students. Finally, it may be worthwhile for teachers and students to design the learning tasks together in order to more accurately determine the needs of students in particular classrooms. The same frameworks and tools used in SLA research can be used to assess learners' needs in the classroom, as well as being used in instruction and practice.

REFERENCES

*ALAMARI, Manal. 1982. **Type and Immediate Effect of Error Correction for Hebrew Second-language Learners.** Jerusalem: School of Education, Hebrew University. (In Hebrew)*

*ALLWRIGHT, R.L. 1975. **Problems in the Study of the Language Teacher's Treatment of Learner Error.** In M.K. Burt & H.C. Dulay (eds.), New Directions in Second Language Learning, Teaching and Bilingual Education. Washington, D.C.: Teachers of English to Speakers of Other Languages (TESOL), 96-109.*

*BEEBE, L., & Giles, H. 1984. **Speech Accommodation Theories:** A Discussion in Terms of Second Language Acquisition. International Journal of the Sociology of Language, 46, 5-32.*

*BERNHARDT, E. 1991. **Reading Development in a Second Language.** Norwood, NJ: Ablex.*

*BERRY, C., & Bailey, D. 1983. **Video English.** London: MacMillan.*

BILLMYER, Kristine. 1990. "I Really Like Your Lifestyle": ESL Learners Learning how to Compliment. Penn Working Papers in Education Linguistics, 6 (2), 31-48.

BLUNDELL, J., Higgens, H., & Middlemass, N. 1982. Functions in English, Oxford: Oxford University Press.

BROWN, H.D. 1991. Breaking the Language Barrier. Yarmouth, ME: Intercultural Press

CANALE, M., & Swain, M. 1980. Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing. Applied Linguistics, 1, 1-47.

CARROLL, S., Roberge, Y., & Swain, M. 1991. The Role of Feedback in Adult Second Language Acquisition: Error Correction and Morphological Generalization. Toronto: Modern Language Centre, Ontario Institute for Studies in Education.

CARROLL, S., & Swain, M. 1991a. Explicit and Implicit Negative Feedback: An Empirical Study of the Learning of Linguistic Generalizations. Toronto: Modern Language Centre, Ontario Institute for Studies in Education.

CARROLL, S., & Swain, M. 1991b. The Role of Negative Feedback in Adult Second Language Learning. Toronto: Modern Language Centre, Ontario Institute for Studies in Education.

CHAUDRON, C. 1988. Second Language Classrooms: Research on Teaching and Learning. Cambridge: Cambridge University Press.

COHEN, A.D. 1990. Language Learning: Insights for Learners, Teachers, and Researchers. NY: Newbury House/Harper Collins.

COHEN, A.D., & Olshtain, E. 1981. Developing a Measure of Sociocultural Competence: The Case of Apology. Language Learning, 31, 113-134.

COHEN, A.D., Olshtain, E., & Rosenstein, D.S. 1986. Advanced EFL Apologies: What Remains to be Learned. International Journal of the Sociology of Language, 62, 51-74.

DICKERSON, L. 1975. The Learner's Interlanguage as a Set of Variable Rules. TESOL Quarterly, 9, 401-408.

ELLIS, R. 1989. Are Classroom and Naturalistic Acquisition the same? Studies in Second Language Acquisition, 11 (3), 305-328.

ELLIS, R. 1990. **Instructed Second Language Acquisition.** Oxford: Basil Blackwell.

HEATH, S.B. 1986. **Children at Risk? Building Investments in Diversity.** Stanford, CA: School of Education, Stanford University.

HOLLEY, F.M., & King, J.K. 1971. **Imitation and Correction in Foreign Language Learning. Modern Language Journal,** 55 (8), 494-498.

KRASHEN, S. 1982. **Principles and Practice in Second Language Acquisition.** Oxford: Pergamon.

KRASHEN, S.D. 1985. **The Input Hypothesis:** Issues and Implications. London: Longman.

KRASHEN, S.D. 1989. **Comprehensible Input and Some Competing Hypotheses.** Los Angeles: Department of Linguistics, University of Southern California.

LARSEN-FREEMAN, D. 1983. **Training Teachers or Educating a Teacher.** In J. Alatis, H.H. Stern, & P. Strevens (eds.), Georgetown University Round Table on Languages and Linguistics 1983. Washington, D.C.: Georgetown University Press, 264-274.

LARSEN-FREEMAN, D. 1990. **On the Need for a Theory of Language Teaching.** In J. Alatis (ed.), Languages and Linguistics 1990. Washington, D.C.: Georgetown University Press, 261-270.

LARSEN-FREEMAN, D. 1991. **Teaching Grammar.** In M. Celce-Murcia (ed.), Teaching English as a Second or Foreign Language. 2nd ed. NY: Newbury House/Harper Collins, 279-296.

LARSEN-FREEMAN, D., & Long, M. 1991. **An Introduction to Second Language Acquisition Research.** London: Longman.

LIGHTBOWN, P.M. 1985. **Great Expectations:** Second-language Acquisition Research and Classroom Teaching. Applied Linguistics, 6 (2), 173-189.

LONG, M.H. 1983. **Natives Speaker/non-native Speaker Conversations in the Second Language Classroom.** In M.A. Clarke & J. Handscombe (eds.), On TESOL '82: Pacific perspectives on Language Learning and Teaching. Washington, D.C.: TESOL, 207-225.

LONG, M.H. 1988. **Instructed Interlanguage Development.** In L.M. Beebe (ed.), Issues in Second Language Acquisition. NY: Newbury House/Harper & Row, 115-141.

LOVEDAY, L. 1982. **The Sociolinguistics of Learning and Using a Non-native Language.** Oxford: Pergamon.

*MUNBY, J. 1979. **Teaching Intensive Reading Skills.** In R. Mackay, B. Barkman, & R.R. Jordan (eds.), Reading in a Second Language. Rowley, MA: Newbury House, 142-158.*

*NATION, I.S.P. 1990. **Teaching and Learning Vocabulary.** NY: Newbury House.*

*NEUFELD, G.G. 1977. **Language Learning Ability in Adults: A Study of the Acquisition of Prosodic and Articulatory Features.** Working Papers on Bilingualism, 12, 45-60. ERIC ED 135 233.*

*NEVO, N. 1989. **Test-taking Strategies on a Multiple-choice Test of Reading Comprehension.** Language Testing, 6 (2), 199-215.*

*OLSHTAIN, E., & Cohen, A.D. 1989. **Speech Act Behavior Across Languages.** In H.W. Dechert et al. (eds.), Transfer in Production. Norwood, NJ: Ablex, 53-68.*

*OLSHTAIN, E., & Cohen, A.D. 1990. **The Learning of Complex Speech Act Behavior.** TESL Canada Journal, 7 (2), 45-65.*

*PICA, T. 1989. **Research on Language Learning: How Can It Respond to Classroom Concerns?** Penn Working Papers in Educational Linguistics, 5 (2), 1-27.*

*RILEY, P. 1981. **Towards a Contrastive Pragmalinguistics.** In J. Fisiak (ed.), Contrastive Linguistics and the Language Teacher. Oxford: Pergamon, 121-146.*

*ROSENSTEIN, D.S. 1982. **The Efficacy of Correcting Recurrent Oral Errors among E.F.L. Students.** Beersheva, Israel: Department of English as a Foreign Language, Ben-Gurion University.*

*RUBIN, J. 1975. **What the 'Good Language Learner' can Teach us.** TESOL Quarterly, 9, 41-51.*

*SCHACHTER, J. 1983. **A New Account of Language Transfer.** In S. Gass & L. Selinker (eds.), Language Transfer in Language Learning. Rowley, MA: Newbury House, 98-111.*

*SCHACHTER, J. 1984. **A Universal Input Condition.** In W.E. Rutherford (ed.), Language Universals and Second Language Acquisition. Philadelphia: John Benjamins, 167-183.*

*SCHLUE, K. 1977. **An Inside View of Interlanguage.** In C.A. Henning (ed.), Proceedings of the Los Angeles Second Language Research Forum. Los Angeles: ESL Section, Department of English, University of California, 342-348.*

*SCHMIDT, R.W., & Richards, J.C. 1981. **Speech Acts and Second Language Learning.** Applied Linguistics, 1 (2), 129-157.*

SCHMIDT, R.W., & Frota, S.N. 1986. **Developing Basic Conversational Ability in a Second Language: A Case Study of an Adult Learner of Portuguese.** In R.R. Day (ed.), *Talking to Learn: Conversation in Second Language Acquisition.* Rowley, MA: Newbury House, 237-326.

SWAN, M., & Walter, E. 1985. **Cambridge English Course.** Book 2, Cambridge: Cambridge University Press.

TARONE, E. 1988. **Variation in Interlanguage.** London: Edward Arnold.

TARONE, E., & Yule, G.. 1989. **Focus on the Language Learner: Approaches to Identifying and Meeting the Needs of Second Language Learners.** Oxford: Oxford University Press.

TOMASELLO, M., & Herron, C. 1989. **Feedback for Language Transfer Errors.** Studies in Second Language Acquisition, 11 (4), 385-395.

VanPATTEN, B. 1990. **Attending to Form and Content in the Input: An Experiment in Consciousness.** Studies in Second Language Acquisition, 12 (3), 287-301.

WENDEN, A. 1985. **Learner Strategies.** TESOL Newsletter, 19 (1), 4-5, 7.

WOLFSON, N. 1989. **Perspectives: Sociolinguistics and TESOL.** Cambridge: Newbury House.

WOLFSON, N., & Judd, E. 1983. **Sociolinguistics and Language Acquisition.** Rowley, MA: Newbury House.

YULE, G., & Macdonald, D. 1990. **Resolving Referential Conflicts in L2 Interaction: The Effect of Proficiency and Interactive Role.** Language Learning, 40 (4), 539-556.

YULE, G., Yanz, J., & Tsuda, A. 1985. **Investigating Aspects of the Language Learner's Confidence: An Application of the Theory of Signal Detection.** Language Learning, 35, 473-88.

YULE, G., Powers, M., & Macdonald, D. 1991. **The Variable Effects of Some Task-based Learning Procedures on L2 Communicative Effectiveness.** Baton Rouge, LA 70803: Linguistics Program, Louisiana State University.

FOOTNOTES

1. Revised version of a Paper prepared for presentation at the RELC Regional Seminar on "Language Acquisition and the Second/Foreign Language Classroom," 22-26 April 1991. We wish to thank Jim Coady for helpful comments regarding L2 vocabulary acquisition and reading studies.

2. Krashen (Personal Communication) and others would still argue that pure comprehensible input is the fastest means of acquiring a language and that "tampering" with it by teaching language forms may disturb the nature process, so the debate continues.

3. In three experiments involving the teaching of a rule to adult ESL and French L2 learners, with four feedback conditions and assessment of short- and long-term learning, Suzanne Carroll and Merrill Swain found negative feedback to be useful, depending on the rule being learned and the level of the learner (Carroll & Swain 1991a, 1991b; Carroll, Roberge, & Swain 1991).

4. Eskimo had to be dropped from the experiment because only two judges could be found, who spoke different dialects of Eskimo and who frequently disagreed greatly in their ratings.

5. The major semantic formulas, any one or combination of which would suffice to represent the particular speech act. Thus, offering repair ("Let me pick that up!") could serve as an apology, or expression of apology plus acknowledgement of responsibility ("I am sorry. I didn't see you.").

6. Of Course, students could also be asked to indicate why they chose a given multiple-choice alternative (Munby 1979) and/or to indicate the strategies that they used in selecting that alternative (Nevo 1989).

7. At the outset it was indicated that this paper would touch on some of the areas in which SLA research has contributed to classroom instruction. Two major areas which were not addressed in this paper but which are worthy of mention are those of vocabulary acquisition and L2 reading. Nation (1990), for example, offers an extensive review of literature on L2 vocabulary acquisition and then provided suggestions based on those studies as to the teaching of high and low frequency words and paired opposites, among other things. Bernhardt (1991) surveys the research on L2 reading and supports the view of reading as an interaction between the text and the reader, and would advocate the use of authentic texts.

THE MATCHED-GUISE TECHNIQUE FOR MEASURING ATTITUDES AND THEIR IMPLICATIONS FOR LANGUAGE EDUCATION: A CRITICAL ASSESSMENT

Stephen J. Gaies and Jacqueline D. Beebe

Many groundbreaking studies on language attitudes and their relevance in educational settings have made use of subjective reaction tests. One of the most prominent elicitation devices has been the matched-guise technique. Although this technique has great intuitive appeal by virtue of its ability to control may be more illusory than real. Furthermore, the validity of this measure of subjective reactions to linguistic codes may depend on the nature of the response that subjects make to the guises. The purpose of this paper is to offer a critical assessment of the methodology of the matched-guise technique and to suggest how it might be best employed in future research on the manifestations and consequences, in educational settings, of attitudes toward particular linguistic codes. Some recent efforts to apply the matched-guise technique to investigate different aspects of English language teaching in Japan are described.

Attitudes and Research on Language Acquisition and Use

There has been almost universal recognition among teachers researchers that attitudes are a key variable in language development, whether instructed or noninstructed, in a second-language or a foreign-language setting, among children and adults alike. To put it simply, we understand that the attitude that people have toward a language and toward members of a language community have a lot to do with whether and how those people develop proficiency in that language. This understanding is very clearly reflected in some theories of second language development -- the Acculturation Model (Schumann, 1978) and Accommodation Theory (Beebe & Zuengler, 1983) are probably the two best known efforts to assign a central role to attitudinal variables in second language acquisition -- but even the more mentalist theories of second language development often explicitly recognize the importance of attitudes -- for example, in the "affective filter" in Krashen's (1985) Input Hypothesis.

There has, of course, been considerable debate about what constitute reliable and valid measures of attitudes. However, unlike many other constructs of affect or personality, there is evidence that at least some attitudes related to the perception of group characteristics -- for example, stereotypes of outgroups -- are highly stable and that measures of these attitudes can be highly reliable (see, for example, Oller, 1979).

It also appears that subjective evaluations of social dialects or foreign languages are quite uniform throughout a speech community and that these community "norms" are firmly established as early as school-entering age (Rosenthal, 1974). In addition, we know that in any given speech community, evaluations of speech are systematically related to the presence of particular linguistic variants; the work of Labov (1972a,b) has been particularly influential in this area.

What is the Matched-Guise Technique?

The matched-guise technique is the use of recorded voices of people speaking first in one dialect or language and then in another; that is, in two "guises". ... The recordings are played to listeners who do not know that the two samples of speech are from the same person and who judge the two guises of the same speaker as though they were judging two separate speakers. (Richards, Platt, & Weber, 1985, p. 171)

The matched-guides technique has two basic purposes:

(a) to elicit reactions to particular codes by having subjects respond to taped samples of those codes, rather than by having subjects express opinions about the codes themselves, and

(b) to control all variables other than the codes themselves.

The hope is that subjects will assume that each sample has been produced by a different speaker (and not by the same bilingual or bidialectal individuals). Thus, in a typical application of the matched-guise technique, we might wish to investigate whether members of a bilingual community have different attitudes to the languages of that community by having bilingual speakers produce languages or dialects is the same, we interpret any differences in subjects' reactions to the two

sets of samples to indicate differences in their attitudes toward the linguistic codes and/or the community of speakers of that code.

The matched-guise technique has been applied to a very broad range of sociolinguistic, social psychological, and educational issues:

a. attitudes of foreign-language learners toward target-language speakers and the target-language community (culture) (e.g., Gardner & Lambert, 1972;)

b. the linguistic bases of teacher prejudice (e.g., Ford, 1984; Williams, 1970, 1973a,b)

c. attitudes toward different varieties (e.g., Fremder, Brown, & Lambert, 1970; Fremder & Lambert, 1973), codes (e.g., Bourhis & Giles, 1976; Tucker & El-Dash, 1975)

d. attitudes toward the speech of language learners or nonnatives (e.g., Fayer & Krasinski, 1987)

e. the phenomena of convergence and divergence

f. the effect of speaker and hearer variables (such as gender and/or perceived expertise) on comprehension, recall, or evaluation (e.g., Markham, 1988)

Furthermore, as Chaika (1989) argues, the matched guise technique and other subjective-reaction tests are of value not only for research purposes, but also in job training. Chaika suggests, for example, that "teachers and social workers who need to realize that they may unconsciously evaluate pupils and clients unfavorably just because of their pitch, loudness, tempo, timber, and intonation" (pp. 57-58). Similarly, Milmoe, Rosenthal, Blane, Chafetz, & Wolf (1967) argue that the matched-guise technique can be used to sensitize doctors (and presumably other professionals) to how their own speech will be interpreted and evaluated by their clients.

This paper reports, in very summary form, two recent applications of the learners of English. Each study is, to our knowledge, the first investigation of its kind in the Japanese EFL context, and each explores a phenomenon that has been the subject of much recent discussion. Finally, the studies serve to illustrate some of the challenges, problems, and limitations involved in using the matched-guise

technique in its original form: issues that we will enumerate and briefly discuss below.

Two Recent Applications of the Matched-Guise Technique to the Investigation of Japanese Learners of English as a Foreign Language

Beebe, Harmon, and Kushibuchi (1990) were interested in the reception extended to students who have lived abroad and have then returned to mainstream high schools with possible deficiencies in their Japanese cultural competence but spoken English abilities surpassing those of their classmates and often of their teachers. They looked for evidence to support or cast doubt on the conventional wisdom that returnees may find their status a social liability in both their EFL classes and the rest of their high school life.

Currently, more than 50,000 Japanese of school-age undergo schooling overseas. In his sociolinguistic survey of contact between Japanese and other languages, Loveday (1986) mentions the educational and social difficulties -- including rejection and ridicule by Japanese classmates -- faced by "returning youngsters" (Nihonjin-kikoku-shijo) -- as they (re)enter the Japanese educational system after their period of residence abroad.

White (1988) maintains that the Japanese educational establishment perceives school-age returnees as suffering from an illness from which they must recover; each sign of deviance in dress, eating habits, liberal use of English loan words, etc., is taken more seriously by teachers when it is exhibited by a returnee than when a "normal" Japanese child does the same thing. White also tells of a returnee who was called gaijin, "foreigner" by her classmates.

Sato (1982) reports on the peer pressure returnees face to "'forget' their 'foreign' English and adapt to Japanese-style English" and mentions a returnee who assumes that her teacher will also be pleased with her linguistic readjustment (pp. 67-68). Beebe has seen a returnee with near-native pronunciation of English unconsciously pronounce English words in conformity to the five vowel, open-syllable sound system of Japanese when she demonstrated how she taught English grammar to secondary school students preparing for entrance examinations at a private coaching school. The returnee explained that she had been instructed by the school management to model her lessons after those of the other Japanese teachers of English.

The Beebe, Harmon, and Kushibuchi study assessed attitudes of female students in four Tokyo high schools toward speakers of either English or Japanese who were introduced as Japanese high school students who either had or had not lived in the United States. The research questions were:

1. Will the mean scores of responses to speakers perceived as returnees differ from the mean scores for the same speakers when perceived as non-returnees?

2. Will the mean scores for speakers speaking English differ from the mean scores form the same speakers speaking Japanese?

The speech samples were unscripted one-minute versions of the story "The Tortoise and the Hare" told by bilingual females. The study was administered by the students' American teacher during their English class. A total of 151 questionnaires, written in Japanese, in which subjects responded to one English and to one Japanese guise, were randomly selected for analysis.

The questionnaire asked each subject to rate the extent to which she believed the speaker would share her own views on three controversial issues and to rate the speaker on nine traits anecdotally ascribed to returnees or representing the three personality dimensions distinguished by Lambert (1967) of competence, personal integrity, and social attractiveness.

Results consistently but only very slightly favored returnees over non-returnees. Only one question "I'd like to be friends with this person," prompted a statistically significant difference; a higher score for the Japanese-speaking returnee guise than for the Japanese-speaking non-returnee guise. Furthermore, the judges perceived themselves as holding views almost as similar to those of the returnees as to those of the non-returnees.

While the variable of residential history exerted little effect on scores, English guises showed both a consistent and strong advantage over Japanese guises. On seven out of nine personality traits and on two out of three shared-view scores, English speakers scored significantly higher. While attractiveness as a friend was substantially higher for the English-speaking guises, the strongest effects were for traits representing competence: leadership, worldliness/open-mindedness, intelligence, and language ability. The attitude toward English in Japan appears to

resemble the pattern found in some diglossic speech communities, in which the use of the high, or prestige code, especially enhances perceptions of competence.

It must of course be recognized that the study examined only the reactions of female subjects to female guises. Under any circumstances, the need to confirm whether the response patterns would hold for comparable male subjects reacting to male guises and for cross-sex comparisons is evident. In the particular instance of Japanese school-age returnees, this issue may be especially important, since White (1988) cites evidence that because parents have more conventional future career goals for sons than for daughters they place more importance on strictly reacculturating returnee males in special readjustment schools, while a greater proportion of returnee females are placed back in mainstream Japanese schools.

The second study employing the matched-guise technique to explore language attitudes and their educational implications in Japan explored simultaneously perceptions of code-switching as a linguistic phenomenon and reactions to code-switchers. Furuya-Nakajima and Vogt (1990) were interested in examining the attitudes of monolingual Japanese studying English conversation in school toward code-switching and to individuals who code switch.

Previous research presents very consistent evidence that

(a) code switching is rule-governed, and not simply the result of random alternations between two codes (see, for example, Pfaff, 1970; Poplack, 1980);

(b) code switching is not the result of a lack of proficiency in one or both of the codes (see, for example, Appel & Muysken, 1987; Poplack, 1980); but

(c) both monolinguals (see Grosjean, 1982) and code switchers themselves (see, for example, Amuda, 1986; Chana & Romaine, 1984) often hold negative attitudes toward code switching and toward those who engage in code switching.

The phenomena explored by previous research on code switching has not examined Japanese/English code switching. Since the position of English in Japan is distinctively different from code-switching situations that have been the focus of previous research, mention of a few aspects of the status of English in Japan may be valuable.

Anyone with even a casual interest in contemporary Japan will have observed or heard about the pervasiveness of English in Japanese public settings: in the media, on T-shirts, in almost any use of language intended for public consumption. According to Haarman (1984), the pervasive use of English in the mass media simply reflects stereotyped conceptions and social values in Japan to foreign languages. On the basis of his survey, Haarman concludes that the impact of English on Japan has created what he calls an "impersonal bilingualism" which is unrelated to any features of monolingual interaction in everyday community life in Japan.

Beneath this veneer of English, however, is a potentially more important phenomenon: the massive amount of English lexicon that has been absorbed into Japanese. Stanlaw (1982), for example, estimates that 8 percent of the contemporary Japanese lexicon is English-based. This figure, however accurate (see, for example, the reservations voiced by Loveday, 1986), reflects the strong amount of contact between English and Japanese. There is evidence, moreover, that at least some elements of this bilingualism, "impersonal" or not, do not elicit a uniform reaction among native speakers of Japanese. Ishino (1983), for example, found that for certain borrowings, different age groups differed by as much as 60 percent in their judgment of the acceptability of the borrowings as Japanese: Whereas the loan for live (which is invariably written in katakana, the script for loanwords, and which refers to a pop or rock concert) was considered as "Japanese" by only 16 out of 100 informants in their 50's and in positions of authority, 79 out of 104 university students judged the loan to be acceptable.

Furuya-Nakajima and Vogt thus hypothesized that the linguistic competence of code-switchers would be evaluated negatively by monolingual speakers of Japanese. They further hypothesized that their subjects would attribute more unflattering personality characteristics --- those of someone "non-Japanese" or "contaminated" -- to the code-switching speakers than to the same speakers using Japanese only.

Two Japanese females in their early 30's, balanced bilinguals who had lived abroad and attended international schools, each gave the directions to her house, once in Japanese only, and once using English/Japanese code-switching.

The subjects who judged the guises were 22 females and 6 males ranging in age from eighteen to twenty-one. They were attending two intermediate-level college English classes, and the study was carried out in Japanese by a Japanese researcher.

From their responses to the open-ended questions that served to elucidate and to validate the ratings on the semantic differential scales, it was clear that these subjects did not view code switching with much admiration and perceived it as an inappropriate response to English speakers who understand some Japanese but not enough to take in a message entirely in Japanese.

Thus, like monolinguals elsewhere, these Japanese subjects exhibit evidence of negative attitudes toward code switching, attitudes that lead them to respond differentially to monolinguals and to code-switchers. Speakers were perceived as "standing out" or "flashy" in their Japanese/English guise; they were also rated as more sociable, more entertaining, more creative, more flexible, more ambitious, and more self-confident; in their Japanese guise, on the other hand, they were judged to be more intelligent, more dependable, more sincere, and more well-mannered.

Methodological Issues in the Use of the Matched-Guise Technique

It must be remembered that the matched-guise technique is simply that: a technique for eliciting attitudes toward language codes and their users. Any research employing the matched-guise technique must of course strive to eliminate, or at least to minimise, threats to internal and external validity. (Campbell & Stanley, 1963, provide the list that is most often drawn upon in discussion of experimental validity.) Such research must also demonstrate that the instruments used provide a reliable measure of the behavior being investigated.

Thus, we wish to limit our discussion to those methodological issues that apply particularly or uniquely to the use of the matched-guise technique to investigate attitudes toward and about second/foreign language education and the settings in which such instruction takes place. We will discuss six such issues (recognizing, of course, that others might, in a more comprehensive review, also be worthy of discussion).

1. The use of matched guises: Is it necessary? Is it sufficient?

The first, and in some ways the crucial issue, concerns the very nature of the matched-guise technique: namely, the question of what precisely is and needs to be "matched." One aspect of this issue that has received considerable attention is the relative advantages having speakers read from scripts (including control of linguistic variables) and of having speakers produce semi- or unscripted speech. An

equally important issue, however, concerns the necessity of controlling variables of voice quality, which is the primary justification for the matched-guise technique. Hudson (1980), for example, has argued that in some cases the use of guises may be counterproductive, since there is a danger that a speaker "may be producing an exaggerated version of the accents or dialects he is simulating" (p. 205). Hudson goes on to argue that since "there is little difference between results produced by the matched guise technique and those where the voices were each produced by a different speaker" (p. 205), the use of the matched-guise technique may be unnecessary.

To argue that the use of bilinguals is unnecessary and that "comparable" monolinguals might be used to prepare speech samples raises other questions:

(a) Is the use of bilinguals sufficient to produce "matched" samples of speech, and

(b) How do we establish comparability (regardless of whether bilinguals or monolinguals are used).

We might begin with the very phenomenon that Hudson cites: the comparability of voice quality. We know that meanings and values are assigned to voice qualities: Streeter et al. (1977), for example, found that deceptive speech tended to have a higher pitch than truthful speech and that higher-pitched speech was judged to be less truthful by listeners. We also know that voice qualities are not valued similarly across speech and language communities: Chaika (1989), for example, discusses the "meanings" communicated by changes in pitch by American English-speaking Blacks.

Thus, one might imagine that the characteristic high pitch that Japanese females appear to cultivate would almost certainly elicit less favorable reactions from American English-speaking subjects than from Japanese-speaking subjects. However, would Japanese subjects respond differently to high- and low-pitched English spoken by females? For this reason, the degree of control involved in the use of the matched-guises are delivered in such a way as to produce voice qualities that are equally valued by (or at least equally typical of) their respective language communities, can we be sure that two guises of the same speaker are any more comparable than the voices of two <u>different</u> speakers?

2. The use of a limited number of speakers: Is there as much variation within groups as between groups?

Many studies employing the matched-guise technique have used one or a very limited number of speakers. This raises the question of how consistently subjects react to different speakers of the same language or, to put it another way, whether reactions to different speakers of the same code vary just as much as do reactions to speakers of different codes. In both of the studies that we have summarized in this paper, there is evidence that characteristics of individual speakers may have had considerable influence on the reactions of the subjects. In the code-switching study, the two speakers interpreted the researchers' directions to speak as if to a friend differently and one speaker employed a more formal register of Japanese than the other. In the returnee study, one of the two speakers provoked the strongest negative reactions in her Japanese guise and the strongest positive reactions in her English guise. This, the researchers speculated, might well have been due to, among other things, reactions to what they perceived to be the highly distinctive "personality cue value" (Webster & Kramer, 1968, p. 239) of the voice of the speaker who elicited the extreme reactions in both languages.

3. How credible (authentic) is the context in which a matched-guise investigation is carried out?

Research on language attitudes has shown that individuals rarely hesitate to form judgments about a speaker's character and personality on the basis of a very limited sample of speech. Nevertheless, the importance of presenting subjects with credible tasks -- that is, tasks that are suited to the subjects' backgrounds and abilities and to the setting in which the research is being carried out -- is worth emphasizing. In the Beebe et al. study, subjects were told that they were to respond to samples of speech produced by high school-age speakers who were being considered for use in recording taped material to accompany a new textbook series. This, we would contend, was a credible task to have learners perform in an English classroom; subjects' reactions would seem to be highly relevant to the selection of speakers to record pedagogical materials.

Often, however, subjects are asked to react to speakers for no particular reason, without specific guidelines, or, worst of all, in a way that forces subjects to pretend that they are someone other than themselves (for examples of this latter problem, see Carranza & Ryan, 1975; Swacker, 1977).

For example, in the Furuya-Nakajima and Vogt study on code switching, the subjects were not told the nationality of the speakers or to whom the directions to the home were being given. The answers to the open-ended questions indicate that the subjects made assumptions of their own and established their own criteria; most assumed that the speaker was a Japanese returnee, and some commented that if the addressee were a foreigner, English should be used, while Japanese ought to be the choice when speaking to a fellow Japanese. (This demonstrates the value of either a written or verbal debriefing of subjects in interpreting the results of judgments.)

The subjects in the Furuya-Nakajima and Vogt study on code switching were in most cases proficient enough in English to understand all of the English used by the code-switching guises; the fact that they offered numerous and strong reactions to the English proficiency of the speakers is interesting, of course, but their judgments may be suspect since they were being asked to compare two unequal alternatives: If a Japanese-English bilingual were speaking to a monolingual Japanese (even one who had studied English in school for as many years as these subjects had), Japanese, rather than Japanese-English code switching, would be the only reasonable alternative. In effect, the subjects were being asked to imagine that they were someone else: a Japanese-English bilingual, or a nonnative speaker of Japanese, for whom a Japanese-only set of directions might have been less understandable or pleasing than directions given in English and Japanese.

4. The need for debriefing: Were subjects properly deceived?

The value of using the matched-guise technique rather than more direct measures (e.g., questionnaires) is that it "appears to reveal judges' more private reactions to the contrasting group" (Lambert, Anisfeld, & Yeni-Komshian, 1965). This is true, however, only if the deception is successful: that is, if subjects actually believed that they were reacting to what they thought were different speakers. This points to the need for either pilot testing (see, for example, Wolck, 1973) or, preferably, debriefing.

The advantage of debriefing is that in any of a number of ways--through open-ended questionnaires, through interviews, or through other means -- subjects can indicate not only whether they were in fact deceived, but also any other reactions that might shed light on the meaning of their response to semantic differential scales (or whatever primary task is involved).

In the Beebe et al. study, some doubt was cast on the believability of the guises in comments by independent listeners who, when questioned about the age of the speakers, judged them to be older than high school age. The study would have been directly questioned about their speculations on the purpose of the study and the identities of the speakers.

Sometimes, the design of a study does not permit debriefing; an example of this is Bourhis and Giles's (1976) investigation of reactions of theatergoers to different versions of a public-address announcement. However, in research on attitudes of language learners, debriefing is normally quite feasible, and it is therefore surprising that it has not been done more regularly.

5. The validity of semantic differential scales.

Semantic differential scales are not the only way to elicit reactions of subjects to samples of speech. There are many other ways to have subjects make judgments about guises, and these, as well as performance tasks have been used in matched-guise research (see Appendix).

However, semantic differential scales have been used so frequently in matched-guise research that they require some comment. As Oller (1979, p. 36) has pointed out, the tendency of [semantic differential] scales to correlate in meaningful ways is about the only evidence we have concerning the validity of such scales. Thus, the fact that negatively valued scales such as "stubborn", "nervous", and "shy" tend to cluster together (by correlation and factor analysis techniques) is the primary basis on which the validity of semantic differential scales is argued.

The problem, however, is that what is positively or negatively valued by a group cannot always be known in advance (see Yamamoto & Swan, 1989, for an empirical investigation of this issue). This is a problem that has been acknowledged since the matched-guise technique was first used (Gardner and Lambert, 1972) and which the returnee study again illustrates. Beebe, Harmon, and Kushibuchi (1990) originally assumed that the Japanese word sukeban, which they translated into English as "rebelliousness" and which in Japanese refers to "bad" teenaged-girls who are apt to be gang leaders, would be a negatively valued trait. But this trait scored much higher for the guise that received positive ratings on most other traits and much lower for the guise that received the lowest ratings overall. This the researchers interpreted to mean that "a dash of rebelliousness was seen [by their subjects] as a positive trait."

In their discussion of their code-switching study, Furuya-Nakajima and Vogt maintain that "ambition" and "self-confidence" are more highly valued in the West than in Japan, and thus the high ratings the code-switching guises received on these two traits could be seen as reflecting a negative evaluation.

The second issue involves differences in meaning between labels, in different languages, for the ends of semantic differential scales. Since it is often difficult to find exact or even close equivalents (compare, for example, the English word friendly with the Spanish word simpatico), there will inevitably be some imprecision in comparing responses made by monolingual speakers of different languages to different guises.

6. To what extent do reactions to speech samples in a matched-guise study correlate with (and perhaps predict) actual overt behavior?

The unresolved questions surrounding semantic differential scales are part of a larger issue: namely, degree to which reactions to speech samples in matched-guise research correlates with (and predicts) actual behavior.

This problem is pervasive in research on attitudes, and only rarely have efforts been made to attempt to investigate the relationship between subjective reactions and actual behavior: for example, Fishman's (1968) use of what he questionnaire about attitudes of Puerto Ricans toward their own ethnicity and attendance at a Puerto Rican cultural evening (to which all subjects had been invited).

In the absence of meaningful correlations between the results of matched-guise research and actual behavior, it becomes difficult to know what one has in fact measured. For example, the surprisingly positive attitudes toward returnees that subjects revealed indirectly in the Beebe et al. study suggest that English ability is a social advantage, but virtually everything that has been reported anecdotally and in a growing body of research into the issue of school-age returnees to Japan suggests that the English ability of returnees is viewed with suspicion.

The Beebe et al. study did not investigate how these subjects would react if they were directly addressed by a peer in fluent English. Would they converge upward toward the fluent speaker of English, or would they become so flustered that the returnee would have to accommodate downward to their ability level?

What we must remember is that the matched-guise technique was designed to elicit stereotypical attitudes toward groups. It may well be such attitudes determine, at least partially, how individuals actually respond to or interact with individual members of their own or some other group, but what is not known, for example, with regard to returnees, is whether the attitudes toward returnees that are tapped by the matched-guise technique correlate with behavior of Japanese students toward returnee classmates.

Conclusion

Insofar as the matched-guise technique itself is concerned, the research that we have reported here suggests what also seems to be true of much previous research: namely, that some of the advantages of the technique as it was originally designed may be more apparent than real; furthermore, there appear some distinct advantages to and no compelling arguments against the use of an alternative approach to the preparation of speech samples: Rather than having the same speaker(s) perform in two or more different guises, several speakers (not necessarily bilinguals) of one code are "matched" subjectively with several speakers of another code, in an effort to produce equivalent sets of samples (see, for example, Alford & Strother (1990).

In conclusion, we research employing the matched-guise technique depend only partially on the technique itself and as much or more on the overall methodological soundness of the research. Methodological considerations aside, however, we would agree with Edwards (1982) that the matched-guise technique provides useful information which can and needs to be confirmed by other means.

THE AUTHORS

Stephen J. Gaies is a professor of English at the University of Northern Iowa. During the 1990-1991 academic year, he is a visiting professor at Temple University Japan. He is the author of Peer Involvement in Language Learning (Harcourt Brace Jovanovich, 1985) and the former editor of the TESOL Quarterly.

Jacqueline D. Beebe is a lecturer in English at the College of Commerce of Nihon University.

REFERENCES

AGHEYISI, R., & Fishman, J. 1970. ***Language Attitude Studies:*** *A Brief Survey on Methodological Approaches. Anthropological Linguistics, 12(5), 137-157.*

ALFORD, R.L., & Strother, J.M. 1990. ***Attitudes of Native and Non-native Speakers Toward Selected Regional Accents of U.S. English.*** *TESOL Quarterly, 24, 479-495.*

d'ANGLEJAN, A., & Tucker, G.R. 1973. ***Sociolinguistic Correlates of Speech Style in Quebec.*** *In Fasold, R.W. (ed.), Variation in the Form and Use of Language: A Sociolinguistics Reader (pp. 327-353). Washington, DC: Georgetown University Press.*

ANISFELD, M., & Lambert, W.E. 1961. ***Social and Psychological Variables in Learning Hebrew.*** *Journal of Abnormal and Social Psychology, 63, 524-529.*

ANISFELD, M., Bogo, N., & Lambert, W.E. 1962. ***Evaluational Reactions to Accented English Speech.*** *Journal of Abnormal and Social Psychology, 65, 223-231.*

APPEL, R., & Muysken, P. 1987. ***Language Contact and Bilingualism.*** *London: Edward Arnold.*

BEEBE, J.D., Harmon, J., & Kushibuchi, T. 1990. ***Attitudes of Japanese High School Students Towards Returnees.*** *Unpublished Manuscript.*

BEEBE, J.D., & Zuengler, J. 1983. ***Accommodation Theory:*** *An Explanation for Style Shifting in Second Language Dialects. In Wolfson, N., and Judd, E. (eds.), Sociolinguistics and Second Language Acquisition (pp. 195-213). Rowley, Mass.: Newbury House.*

BLOM J-P., & Gumperz, J.J. 1972. ***Social Meaning in Linguistic Structures:*** *Code-switching in Norway. In J.J. Gumperz & D. Hymes (eds.), Directions in Sociolinguistics (pp. 404-434). New York: Holt, Rinehart & Winston.*

BOGARDUS, E.S. 1925. ***Measuring Social Distance.*** *Journal of Applied Sociology, 9, 299-308.*

BOUCHARD-RYAN, E. 1973. ***Subjective Reaction Toward Accented Speech.*** *In Shuy, R. W., and Fashold, R.W. (eds.), Language Attitudes: Current Trends and Prospects (pp.60-73). Washington, DC: Georgetown University Press.*

BOURHIS, R., & Giles, H. 1976. ***The Language of Cooperation in Wales:*** *A Field Study. Languages Sciences, 42, 13-16.*

BOURHIS, R. Giles, H., & Lambert, W. 1975. ***Social Conquences of Accommodating One's Style of Speech:*** *A Cross-national Investigation. International Journal of the Sociology of Language, 6, 55-71.*

CAIRNS, E., & Duriez, B. 1975. ***The Influence of Speakers' Accent on Recall by Catholic and Protestant School Children in Northern Ireland.*** *British Journal of Social and Clinical Psychology, 15, 441-442.*

CAMPBELL, D.T., & Stanley, J.C. 1963. ***Experimental and Quasi-experimental Designs for Research on Teaching.*** *In N.L. Gage (ed.), Handbook of Research on Teaching (pp. 171-246). Chicago: Rand McNally.*

CARRANZA, M., & Ryan, E.B. 1975. ***Evaluative Reactions of Bilingual Anglo and Mexican American Addolescents Towards Speakers of English and Spanish.*** *International Journal of the Sociology of Language, 6, 83-104.*

CHAIKA, E. 1989. ***Language:*** *The Social Mirror (2nd ed.). Cambridge, Mass.: Newbury House.*

CHANA, U., & Romaine, S. 1984. ***Evaluative Reactions to Panjabi/English Code-switching.*** *Journal of Multilingual and Multicultural Development, 5, 447-473.*

CHIHARA, R. & Oller, J.W., Jr. 1978. ***Attitudes and Attained Proficiency in EFL:*** *A Sociolinguistic Study of Adult Japanese Speakers. Language Learning, 28, 55-68.*

COOPER, R.L. & Fishman, J.A. 1974. ***The Study of Language Attitudes.*** *International Journal of the Sociology of Language, 3, 5-19.*

COOPER, R.L., Fishman, J.A., Lown, L., Scheier, B., & Seckback, F. 1977. ***Language, Technology and Persuasion in the Middle East:*** *Three Experimental Studies. In Giles, H. (ed.), Language, Ethnicity and Intergroup Relations (pp. 83-98). London: Academic Press.*

DOWNS, L., & Olson, T. 1989. ***What Happens to Japanese Who Graduate from ASIJ?*** *ASIJ Shimbun.*

EL-DASH, L., & Tucker, G.R. 1975. ***Subjective Reactions to Various Speech Styles in Egypt.*** *International Journal of Language, 6, 33-54.*

FASOLD, R. (Ed.). 1973. ***Variation in the Form and Use of Language:*** *A Sociolinguistics Reader. Washington, DC: Georgetown University Press.*

FASOLD, R. 1984. **The Sociolinguistics of Society.** Oxford: Basil Blackwell.

FAYER, J.M., & Krasinski, E. 1987. **Native and Non-native Judgments of Intelligibility and Irritation.** Language Learning, 37, 313-326.

FISHMAN, J.A. 1968. **Readings in the Sociology of Language.** The Hague: Mouton.

FORD, C.E. 1984. **The Influence of Speech Variety on Teachers' Evaluation of Students with Comparable Academic Ability.** TESOL Quarterly, 18, 25-40.

FREMDER, R., & Lambert, W.E. 1973. **Speech Style and Scholastic success: The Tentative Relationships and Possible Implications for Lower Class Children.** In Shuy, R. (ed.), Sociolinguistics: Current Trends and Prospects (pp. 237-271). Washington, DC: Georgetown University Press.

FREMDER, R., Brown, B., & Lambert, W.E. 1970. **The Role of Speech Characteristics in Scholastic Success.** Canadian Journal of Behavioral Sciences, 2, 299-306.

FURUYA-NAKAJIMA, N., & Vogt, L.D. 1990. **"Nanka I Think Speaking 'Chanpon' is Ichiban Natural": A matched-guise Experiment.** Unpublished Manuscript.

GAERTNER, S.L., & Bickman, L. 1971. **Effects of Race on the Elicitation of Helping Behaviour: The Wrong Number Technique.** Journal of Personality and Social Psychology, 20, 218-222.

GARDNER, R.C., & Lambert, W.E. 1972. **Attitudes and Motivation in Second Language-Learning.** Rowley, Mass.: Newbury House.

GENESEE, F., & Bourhis, R.Y. 1988. **Evaluative Reactions to Language Choice Strategies: The Role of Sociocultural Factors.** Language and Communication, 8, 229-250.

GILES, H., Baker, S., & Fielding, G. 1975. **Communication Length as a Behavioural Index of Accent Prejudice.** International Journal of the Sociology of Language, 6, 73-81.

GILES, H., & Powesland, P. 1975. **Speech Style and Social Evaluation.** London: Academic Press.

GILES, H., & Robinson, W.P. (Eds.). 1990. **Handbook of Language and Social Psychology.** New York: John Wiley & Sons.

HAARMAN, H. 1984. ***The Role of Ethnocultural Stereotypes and Foreign Languages in Japanese Commercials.*** *International Journal of the Sociology of Language, 50, 101-121.*

HOFFER, B. 1990. ***English Loanwords in Japanese:*** *Some Cultural Implications. Language Sciences, 12(1), 1-21.*

HOOVER, M.R. 1978. ***Community Attitudes Toward Black English.*** *Language in Society, 7, 65-87.*

HOPPER, R. 1977. ***Language Attitudes in the Job Interview.*** *Communication Monographs, 44, 346-351.*

HUDSON, R.A. 1980. ***Sociolinguistics.*** *Cambridge: Cambridge University Press.*

ISHINO, H. 1983. ***Gendai-gairaigo-ko [Thoughts on Contemporary Foreign Borrowings].*** *Tokyo: Taishukan.*

KIMPLE, J., Jr. 1968. ***Language Shift and the Interpretation of Conversations.*** *Lingua, 23, 127-134.*

KRASHEN, S. 1985. ***The Input Hypothesis.*** *London: Longman.*

LABOV, W. 1972a. ***On the Mechanism of Linguistic Change.*** *In Gumperz, J., & Hymes, D. (eds.), Directions in Sociolinguistics: The Ethnography of Communication (pp. 512-538). New York: Holt, Rinehart and Winston.*

LABOV, W. 1972b. ***Sociolinguistic Patterns.*** *Philadelphia: University of Pennsylvania Press.*

LAMBERT, W.E. 1967. ***A Social Psychology of Bilingualism.*** *Journal of Social Issues, 23, 21-109.*

LAMBERT, W.E., Anisfeld, M., & Yeni-Komshian, G. 1965. ***Evaluational Reactions of Jewish and Arab Adolescents to Dialect and Language Variations.*** *Journal of Personality and Social Psychology, 2, 84-90.*

LAMBERT, W.E., Hodgson, R.C., Gardner, R.C., & Fillenbaum, S. 1960. ***Evaluational Reactions to Spoken Languages.*** *Journal of Abnormal and Social Psychology, 60(1), 44-61.*

LAMBERT, W.E., Frankel, H., & Tucker, G.R. 1966. ***Judging Personality through Speech:*** *A French-Canadian Example. The Journal of Communication, 16, 305-321.*

LOVEDAY, L. 1986. ***Explorations in Japanese Sociolinguistics.*** *Amsterdam: John Benjamins.*

MARKHAM, P.L. 1988. ***Gender and the Perceived Expertness of the Speaker as Factors in ESL Listening Recall.*** *TESOL Quarterly, 22, 397-406.*

MEYER, C. 1988. ***A Voice from Japan:*** *An Outsider Looks In New York: Harcourt Brace Jovanovich.*

MILMOE, S., Rosenthan, R., Blane, H.T., Chafetz, M.E., & Wolf, I. 1967. ***The Doctor's Voice:*** *Postdictor of Successful Referral of Alcoholic Patients. Journal of Abnormal Psychology, 72, 78-84.*

OKAMURA-BICHARD, F. 1985. ***Mother Tongue Maintenance and Second Language Learning:*** *A Case of Japanese Children. Language Learning, 35, 63-89.*

OLLER, J.W., Jr. 1979. ***Language Tests at School:*** *A Pragmatic Approach. London: Longman.*

PASSIN, H. 1980. ***Japanese and the Japanese.*** *Tokyo: Kinseido.*

PENALOSA, F. 1981. ***Introduction to the Sociology of Language.*** *Cambridge, Mass.: Newbury House.*

PFAFF, C. 1979. ***Constraints on Language Mixing:*** *Intrasentential Code-switching and Borrowing in Spanish/English. Language, 55, 291-318.*

POPLACK, S. 1980. ***"Sometimes I'll Start a Sentence in Spanish Y Termino En Espanol":*** *Toward a Typology of Code-switching. Linguistics, 18, 581-618.*

PRESTON, M.S,. 1963. ***Evaluation Reactions to English Canadian French and European French Voices.*** *Unpublished Doctoral Dissertation, McGill University.*

RAMIREZ, A.G., & Milk, R.G. 1986. Notions of Grammaticality Among Teachers of Bilingual Pupils. TESOL Quarterly, 20, 495-513.

RICHARDS, J.C., Platt, J., & Weber, H. 1985. Dictionary of Applied Linguistics. London: Longman.

ROSENTHAL, M. 1974, Spring/Fall. ***The Magic Boxes:*** *Pre-school Children's Attitudes Towards Black Standard English. Florida FL Reporter, 55-93.*

RYAN, E.B., & Giles, H. (Eds.). 1982. ***Attitudes Towards Language Variation:*** *Social and Applied Contexts. London: Edward Arnold.*

SATO, Y. 1982. **Facing Silent Society:** *The Maladjustment of Japanese Returning Children. Harvard Graduate School of Education, pp. 29-.*

SCHUMANN, J. 1978. **The Pidginization Process:** *A Model for Second Language Acquisition. Rowley, Mass.: Newbury House.*

SELIGMAN, C.R., Tucker, G.R., & Lambert, W.E. 1972. **The Effects of Speech Style and Other Attributes on Teachers' Attitudes Toward Pupils.** *Language in Society, 1, 131-142.*

SHAW, M.E., & Wright, J.M. 1967. **Scales for the Measurement of Attitudes.** *New York: McGraw Hill.*

SPOLSKY, B. 1969. **Attitudinal Aspects of Second Language Learning.** *Language Learning, 19, 271-283.*

STANLAW, J. 1982. **English in Japanese Communicative Strategies.** *In Kachru, B.B. (ed.), The Other Tongue: English Across Cultures, (pp. 168-197). Urbana: University of Illinois Press.*

STREETER, L.A., Krauss, R.M., Geller, V., Olson, C., & Apple, W. 1977. **Pitch Changes during Attempted Deception.** *Journal of Personality and Social Psychology, 35, 345-350.*

SWACKER, M.E. 1977. **Attitudes of Native and Non-native Speakers Toward Varieties of American English.** *Unpublished Doctoral Dissertation, Texas A&M University.*

TRUDGILL, P., & Tzavaras, G. 1977. **Why Albanian-Greeks are not Albanians:** *Language Shift in Attica and Biotia. In Giles, H (ed.), Language, Ethnicity, and Intergroup Relations, (pp. 171-184). London: Academic Press.*

WEBSTER, W.G., & Kramer, E. 1968. **Attitudes and Evaluational Reactions to Accented English Speech.** *Journal of Social Psychology, 75, 231-240.*

WHITE, M. 1988. **The Japanese Overseas:** *Can they go home again?. New York: The Free Press.*

WILLIAMS, F. 1970. **Psychological Correlates of Speech Characteristics:** *On Sounding "disadvantaged". Journal of Speech and Hearing Research, 13, 472-488.*

WILLIAMS, F. 1973a. **Some Recent Studies of Language Attitudes.** *In Shuy, R.W. (ed.), Some New Directions in Linguistics, (pp. 121-149). Washington, DC: Georgetown University Press.*

WILLIAMS, F. 1973b. ***Some Research Notes on Dialect Attitudes and Stereotypes.*** *In Fasold, R.W. (ed.), Variation in the Form and Use of Language: A Sociolinguistics Reader, (pp. 354-369). Washington, DC: Georgetown University Press.*

WOLCK, W. 1973. ***Attitudes Toward Spanish and Quechua in Bilingual Peru.*** *In Fasold, R.W. (ed.), Variation in the Form and Use of Language: A Sociolinguistics Reader, (pp. 370-388). Washington, DC: Georgetown University Press.*

YAMAMOTO, M., & Swan, J. 1989. ***Connotative Differences between Foreign and Japanese English Teachers in Japan.*** *Journal of Multilingual and Multicultural Development, 10, 233-253.*

Appendix

Responding to Speech Samples

judgment tasks
- judgments about occupational status
  - guessing a speaker's occupation
  - judging suitability for a specific job/occupation
  - judging employability
- judgments about intellectual/academic potential
  - rating quality of school work
  - predicting future academic success
- forming and changing opinions

performance measures

- **recall tasks** (e.g., Cairns & Duriez, 1976; Markham, 1988)

188

judgments about occupational status

guessing a speaker's occupation (Labov, 1972a) [Sociolinguistic Patterns]
judgments about suitability for a job/occupation
judgments about employability

judgments about intellectual/academic potential

rating quality of school work
predictions about future academic success

helping behavior (Gaertner & Bickman, 1971)

Gaertner and Bickman (1971) had both Black and White callers telephone 540 Black and White subjects, pretending to get the wrong number. Callers told each subject that they were stranded and had used their last dime. Then they requested that the subject call another number to send help. At this number there was a confederate of the caller who recorded which subjects had responded to the caller's request for aid. Black subjects helped Black and White callers equally. Whites helped Blacks less frequently than they helped Whites. This, of course, might also be explained on the basis of racial prejudice; however, it does show that the dialect used does, in and of itself, affect how others treat you. (Chaika, 1989, p. 202)

communication length (e.g., writing letters of recommendation (Giles, Baker, & Fielding, 1975;)

Giles, Baker, and Fielding (1975) developed an experiment to test the reactions of subjects to RP and the nonstandard variety of Birmingham, England. In this application of the matched-guise technique, a male speaker who was proficient in both varieties addressed two groups of Birmingham high school students. The students had to write letters of recommendation stating their opinion of this speaker as a suitable candidate to lecture high school students about the nature of university

studies. The students also had to evaluate him on traditional rating scales. The study was motivated by findings from previous research that subjects write longer letters about someone they like than about someone they don't like. Also, it had previously been demonstrated that subjects speak more when they are conversing with someone they like.

Giles et al. reasoned that if more students wrote letters for the speaker and wrote longer opinions when he was speaking in one guise than in the other, that would demonstrate that the dialect alone caused him to be rated differently. The high school students to whom the speaker had used RP wrote 82 percent more about him than those who heard him talk in his Birmingham accent. Out of 18 subjects, 13 found him "well-spoken" in his RP guise, but only 2 out of 28 subjects who heard him in his Birmingham guise judged him as "well-spoken."

making a decision/choice (e.g., Rosenthal, 1974)

forming and changing opinions (e. g., Giles & Powesland, 1975 [capital punishment]; Cooper, Fishman, Lown, Scheier, & Seckbach, 1977; Cooper & Fishman, 19XX [tobacco and alcohol].

THE INTERACTION HYPOTHESIS: A CRITICAL EVALUATION

Rod Ellis

Introduction

There is now widespread acceptance that the oral interactions in which second language (L2) learners participate provide one of the main sources of data for L2 acquisition. In the case of naturalistic acquisition, the importance of face-to-face interaction with other speakers of the L2 is self-evident. In the case of classroom acquisition, the role of interaction is perhaps less dominant but is nevertheless sufficient for Allwright (1984;156) to call it 'the fundamental fact of pedagogy'. Allwright perhaps exaggerates when he claims that 'everything that happens in the classroom happens through a process of face-to-face interaction' (for, after all, learners often spend a lot of time reading and writing), but it remains true that 'teaching' can be profitably viewed as interaction that supplies learners with opportunities for learning. The study of the relationship between interaction and L2 acquisition, therefore, constitutes one of the main ways in which second language acquisition research (SLA) can inform pedagogy.

The interaction hypothesis provides one theoretical account of this relationship. Other theories also exist (e.g. theories based on the contribution of universal grammar, cf. White, 1990, and theories attributing a more direct role to language production, cf. Pienemann, 1985). However, the interaction hypothesis has received considerable attention during the last decade, it has figured prominently in second language classroom research and it has served as the basis for a number of pedagogical recommendations. For these reasons it warrants careful scrutiny.

The purpose of this paper, then, is to subject the interaction hypothesis to critical evaluation. The paper will begin with an account of the hypothesis, with due attention to the inevitable changes which have taken place since its initial formulation. The next section will consider the evidence that has been cited in support the hypothesis. The section following will attempt an evaluation of the hypothesis in the light of both theoretical arguments and empirical counter evidence. Finally, a revised version of the hypothesis will be proposed.

The Interaction Hypothesis

The interaction hypothesis advances two major claims about the role of interaction in L2 acquisition:

> Comprehensible input is necessary for L2 acquisition.
>
> Modifications to the interactional structure of conversations which take place in the process of negotiating a communication problem help to make input comprehensible to an L2 learner.

The origins of these claims lie partly in the work of Stephen Krashen and partly in that of Evelyn Hatch. Krashen (1977;1980) argued that the subconscious process of 'acquisition' (as opposed to the conscious process of 'learning') occurs when the learner is focussed on meaning and obtains comprehensible input. He emphasised the importance of 'simple codes' (e.g. foreigner and interlanguage talk) and of extralinguistic context for making input comprehensible. He also claimed that language production plays no direct part in 'acquisition'. Hatch (1978) used a 'discourse analysis' approach to study the interactions involving naturalistic child and adult L2 learners and concluded that the regularities which have been shown to exist in the way which learners acquire the grammar of an L2 were the direct result of the kinds of interaction in which they participated. Suggesting that the order of acquisition reflects the differential frequency with which features occur in the input, she argued that the frequency of these features is determined by the efforts which native speakers and learners make to establish and develop a topic through interaction.

The interaction hypothesis itself is most clearly associated with the work of Michael Long. Long (1980) reported on the input and interactional features of native speaker talk to sixteen non-native speakers (all Japanese) in interview-type situations. Input features consist of various purely linguistic aspects of foreigner talk such as breadth of vocabulary used and overall sentence complexity. Interactional features refer to communicative aspects of foreigner talk such as temporal markings and various discourse and topic-incorporation functions (cf. Table 1).[1] Using native-speaker/native-speaker conversations as baseline data, Long discovered that foreigner talk entailed few input modifications but numerous interactional adjustments. In a subsequent article Long (1983) embraced Krashen's views about the role of comprehensible input, arguing that (1) access to it is characteristic of all cases of successful first and second acquisition, (2) greater

Interactional feature	*Definition*	*Example*
Clarification requests	Any expression that elicits clarification of the preceding utterance.	A: She is on welfare. B: What do you mean by welfare?
Confirmation checks	Any expression immediately following the previous speaker's utterance intended to confirm that the utterance was understood or heard correctly.	A: Mexican food have a lot *of* ulcers? B: Mexicans have a lot of ulcers? Because of the food?
Comprehension checks	Any expression designed to establish whether the speaker's own preceding utterance has been understood by the addressee	A: There was no one there. Do you know what I mean?
Self-repetitions:		
(1) repairing	The speaker repeats/ paraphrases some part of her own utterance in order to help the addressee overcome a communication problem	A: Maybe there would be— B: Two? A: Yes, because one *mother* goes to work and the other *mother* stays home.
(2) preventive	The speaker repeats/ paraphrases some part of her own utterance in order to prevent the addressee experiencing a communication problem.	A: Do you share his feelings? Does anyone agree with Gustavo?
(3) reacting	The speaker repeats/ paraphrases some part of one of her previous utterances to help establish or develop the topic of conversation.	A: I think she has *a lot of money*. B: But we don't know that? A: But her husband *is very rich*.
Other-repetitions:		
(1) repairing	The speaker repeats/ paraphrases some part of the other speakers utterance in order to help overcome a communication problem.	A: I think the *fourth family*. B: Not the *fourth family*, the third family.
(2) reacting	The speaker repeats/ paraphrases some part of the other speaker's utterance in order to help establish or develop the topic of conversation.	A: I think *she has three children*. B: This is the thing. *She has three children*.

Source: Based on Pica and Doughty (1985a)

Table 1: Interactional modifications in the negotiation of meaning

(based on Pica and Doughty, 1985)

BEST COPY AVAILABLE

quantities of comprehensible input seem to result in faster acquisition and (3) lack of access to it results in little or no acquisition. Comprehensible input is seen, therefore, as necessary for acquisition, at least for the beginning learner. In the same article, Long argued that modifications to the interactional structure of conversation were 'the most important and widely used' way of making input comprehensible (p.342). He suggests that these are specially facilitative of acquisition because they help to make unfamiliar linguistic input comprehensible. Long does not present any arguments for differentiating the effects of the various interactional modifications, so it must be presumed that it is the quantity rather than quality of the modifications that is important for acquisition. Indeed, as we shall see, the empirical research that has fed off the interactional hypothesis has been based on this assumption.

Long (1983) also considers one of the conditions that promotes the negotiation of meaning. He reports a study (Long, 1980) in which he found a statistically significant higher frequency of various interactional features in NS-NNS as opposed to NS-NS conversations in tasks which required information exchange (i.e. two-way or jigsaw tasks) but not in tasks which did not require any information exchange (i.e. one-way or decision-making tasks). We might add the following to our description of the interaction hypothesis, therefore:

> Tasks in which there is a need for the participants to exchange information with each other promote more interactional restructuring.

It follows that information-exchange tasks also aid comprehension and L2 acquisition.

The interaction hypothesis is also closely associated with the work of Pica. Pica's main contribution has not been in the area of theory construction but in the execution of carefully designed experimental studies designed to test the claims of the interaction hypothesis. As such, it will be considered more fully in the next section. Pica (1987) has, however, extended the interaction hypothesis in one major way. She emphasises the importance of the social relationship between the participants as a determinant of interactional modifications:

> Underlying the need for mutual understanding and the opportunity to modify and restructure social interaction ... is a social relationship in which learners and their interlocutors are aware of their unequal linguistic proficiencies in the second language, but nevertheless see themselves as

> having equivalent status with regard to meeting their needs and fulfilling their obligations as conversational participants. (p.4)

We might, then, add the following to our description of the interaction hypothesis:

> A situation in which the conversational partners share a symmetrical role relationship affords more opportunities for interactional restructuring.

It follows that equality of status between the interactants is also facilitative of comprehension and acquisition.

The interaction hypothesis can now be summarised as hierarchical three-part statement (see Table 2). The first part advances the central claim that learners need to comprehend input in order to develop their interlanguages. The second part states that opportunities to modify the structure of a conversation promotes comprehension. The third part concerns the conditions that create opportunities for restructuring.

---

(1) Comprehensible input is necessary for L2 acquisition (= the input hypothesis).

(2) Modifications to the interactional structure of conversations which take place in the process of negotiating a communication problem help to make input comprehensible to an L2 learner.

(3) a. Tasks in which there is a need for the participants to exchange information with each other promote more interactional restructuring.

b. A situation in which the conversational partners share a symmetrical role relationship affords more opportunities for interactional restructuring.

Table 2: The Interactional Hypothesis

An examination of Long's and Pica's later publications indicates continued adherence to the interaction hypothesis and a preparedness to put forward a number of proposals for pedagogy on the basis of it. Thus Long and Crookes (1987) argue that teachers should make efforts to use fewer display questions because these inhibit the restructuring of interaction that promotes acquisition through comprehensible input. Elsewhere (e.g. Long and Porter, 1985), Long has argued in favour of group work because it promotes greater opportunity for modifying the structure of interactions. Long (1989) has also argued in favour of certain kinds of tasks on the grounds that they produce more and more useful negotiation work. In addition to reaffirming the importance of two-way over one-way tasks, Long also proposes that closed tasks (i.e. a task with a single or a finite set of correct solutions) work better than open tasks. In all these cases, Long refers to the psycholinguistic rationale for his proposals. This rationale is the interactional hypothesis. Pica (1990) also continues to advocate the need for creating the classroom conditions in which the negotiation of meaning can take place, again citing the interactional hypothesis as the theoretical justification.

In two respects, though, Long and Pica's theoretical position does appear to have developed somewhat. First, although Long continues to assert that comprehensible input is necessary for acquisition, he now clearly recognizes that it may not be sufficient (cf. Long, 1989;10). He does not indicate, however, when or in what ways it is insufficient. We are not told whether it is insufficient because its contribution is dependent on some other factor or factors (e.g. Krashen's affective filter) or whether it is insufficient because it cannot account for the whole of acquisition (as White (1987) argues).

Secondly, Long now clearly acknowledges that interaction promotes L2 acquisition not only by supplying comprehensible input but also by providing the learner with opportunities for production. Drawing on the comprehensible output hypothesis[2] (Swain (1985)) he recognises that interlanguage development can take place when learners are 'pushed' to improve their output. In this respect certain interactional modifications may be more helpful than others. For instance, requests for clarification (e.g. 'Pardon') 'stretch' the learner by making her clarify what she has said, whereas confirmation checks do not because they solve the communication problem for the learner. Pica, too, has embraced the output hypothesis. She has been able to demonstrate that although NNSs are less likely to modify their original output than NSs as a result of a communication problem, they do nevertheless modify their own and their interlocutor's productions both semantically and structurally (cf. Pica, 1990, for an excellent survey of her own

work). She also shows that certain types of discourse signals (e.g. requests for clarification) are more likely to promote output modifications. In effect, both Long and Pica are now advancing arguments in favour of the qualitative effects of different types of adjustments. It should be noted, however, that these qualitative effects apply only to production. Where comprehension is concerned it is the sheer quantity of modifications that still appears to count.

The main difference between the early and later work of Long and Pica is that in the latter the interaction hypothesis is no longer seen as the only or even the major explanation of L2 acquisition. It remains however an important element in their psycholinguistic rationale for pedagogic intervention. This paper will focus only on the claims made on behalf of interactional restructuring and comprehension.

A Look at the Evidence

The evidence in support of each part of the interactional hypothesis will be considered separately. It should be noted that because the three parts of the hypothesis are arranged hierarchically, the evidence to be examined is not all of equal weight. To justify the hypothesis it is crucial to demonstrate the first part i.e. to show that comprehensible input is necessary for acquisition (the input hypothesis).

There is, in fact, no direct evidence to support the input hypothesis. Long (1983) comments:

> Like any genuine hypothesis, the input hypothesis has not been proven. There has been no direct test of it to date.

The situation is no different today. It has not even been demonstrated that comprehending messages in an L2 contributes to the acquisition of new linguistic knowledge let alone that comprehension is necessary for its acquisition. A number of studies have reported a relationship between the frequency of features in the input and their acquisition (e.g. Larsen-Freeman, 1976 and Lightbown, 1983) but such studies do not address the input hypothesis as they do not show that the input containing the features was comprehensible to the learners. Chaudron (1988), reviewing studies that have investigated L2 comprehension in a classroom setting, concludes that there is only an 'inkling' of a relationship between comprehensibility and learners' progress. This is an exaggeration; there is no direct evidence of any relationship at all.

The absence of direct evidence, however, does not warrant the rejection of the input hypothesis - as Long (1983) is quick to assert. Indeed, if the weight of indirect evidence is sufficient, the hypothesis can be sustained. Krashen (1985) and Long (1983) point to several kinds of indirect evidence. These are summarised in Table 3. It is difficult to assess this evidence. Some of it does not speak to the input hypothesis at all. For example, there is no study that links foreigner talk to L2 acquisition. Some of it is extremely dubious. For example, the claim that the methods that are supposed to work (e.g. Total Physical Response) do so because they supply more comprehensible input than the methods that are supposed not to work (e.g. audiolingualism) can be disputed both on the grounds that the comparative method studies are seriously flawed (a point Long, but not Krashen, acknowledges) and on the grounds that there is no evidence that the successful methods actually result in more comprehensible input. Some of the evidence is controversial. For instance, researchers do not agree about the role of caretaker talk in L1 acquisition (cf. Gleitman et al,1984). Some of the evidence has also been disputed. Hammerly (1987), for instance, takes a very different view of immersion programs, claiming that they result in pidginization. Even some of the stronger evidence is open to dispute. Some adults, for instance, have proven to be very successful learners even though they had no access to comprehensible input in meaning-focussed communication. It is possible that they worked on the input they received to make it comprehensible (e.g. with the help of a bilingual dictionary) but this involves a very different concept of comprehensible input from that advanced by the input hypothesis. At best the indirect evidence cited provides only weak support in favour of the necessity of comprehensible input. Long's (1983;341) conclusion that 'it is sustained because the predictions it makes are consistent with the available data' is probably not warranted and a safer, more conservative conclusion might be that the hypothesis still awaits confirmation.

Let us now consider the second claim of the hypothesis. What evidence is there to support the view that input obtained from interactional modification is most easily comprehended? Two sets of studies are relevant here - studies which have investigated the effects of input simplifications on comprehension and studies which have investigated whether interactional modifications are especially effective.

Kind of evidence	Brief explanation
Caretaker speech	Caretaker speech to young children is roughly tuned to the children's receptive abilities and is motivated by the need to aid comprehension.
Foreigner talk	Foreigner talk to NSSe is also roughly tuned and functions as an aid to comprehension.
Silent period	Some young children go through a 'silent period' in L2 acquisition. During this period they do not produce but do learn the L2.
Age differences	Older children acquire faster because they obtain more comprehensible input.
Comparative method studies	The studies show that methods that supply plenty of comprehensible input work better than those that supply little.
Immersion programs	Immersion programs have generally been found superior to foreign/ second language programs - again because they supply plenty of comprehensible input. Also, additional exposure to the L2 outside the classroom does not enhance learning, presumably because this does not supply comprehensible input.
Bilingual program	The success of different kinds of bilingual programs reflects the extent to which they supply comprehensible input.
Delayed L1 and L2 acquisition	Studies of children in both L1 and L2 acquisition who are deprived of comprehensible input (e.g. because their parents are deaf) show that acquisition is delayed or non-existent.

Table 3: Indirect evidence in support of the input hypothesis

In general, researchers have concentrated on describing the way in which input is modified in foreigner and teacher talk and have rarely investigated the effects such modifications have on comprehensibility. However, a number of studies indicate that input modifications aid understanding of both spoken and written texts. These studies have been reviewed by Long (1985) and Chaudron (1985;1988). They provide evidence to suggest that (1) rate of speech (e.g. Kelch, 1985), (2) reduction in syntactical complexity (e.g. Johnson, 1981 and Blau, 1982) and (3) increased redundancy (e.g. Chaudron, 1983) result in better comprehension by L2 learners. However, there are problems. First, other studies provide counterfactual evidence. Dahl (1983), for instance, found that L2 learners judgements of the rate of spoken messages did not correlate with the actual speech rate. Speidel et al (1985) found that simple syntax in stories did not improve the reading comprehension of grade 2 pupils in Hawaii. Second, in many of the studies (such as Long 1985) a number of different aspects of input (speech rate, syntax, lexis, discourse) were simplified, making it difficult to determine precisely which aspects contributed to enhanced comprehension.' Chaudron (1983) has also noted that teachers' attempts to simplify may result in too much redundant and confusing information which can inhibit rather than promote comprehension. A safe conclusion is that input modifications probably do facilitate understanding, but precisely which modifications work best and in which combinations remain uncertain.

The studies referred to above, however, are not of direct relevance to the interaction hypothesis, which states that input modifications which derive from attempts to negotiate understanding work best. A study by Pica, Young and Doughty (1986) has addressed this claim directly. For this reason it is an important study and will be considered in some detail. Sixteen low-intermediate learners were divided into two groups. One group received directions requiring them to choose and place items on a small board illustrated with an outdoor scene. These directions were based on those produced in native-speaker/native-speaker interaction but had been systematically premodified with the result that they were longer, more redundant and less complex. The other group received the baseline directions, but were given opportunities to seek verbal assistance if they did not understand. In this way, the researchers aimed to compare the effects of premodified and interactionally adjusted input on comprehension, which they measured as the number of the correct learner responses to the directions. The results showed that interactional modifications did result in higher levels of comprehension. The input derived from the restructured interactions proved to be (1) quantitatively greater, (2) more complex and (3) more redundant than both premodified and baseline input.

Furthermore, a detailed analysis of individual directions showed that 'modifications of interaction were most effective in achieving comprehension when the learners had difficulty in understanding the input but were superfluous when the input was easily understood' (p. 747). this study is often cited as providing strong support for part two of the interaction hypothesis. It is, however, seriously flawed in one major respect. It is impossible to tell whether the advantage shown for the interactionally modified input arose as a result of the greater input which it supplied or, as Pica et al wish to claim, as a result of the opportunities for negotiating meaning. We need to take note of the fact that it was the researchers themselves who determined the quantity of the premodified input and to recognise that we cannot judge whether the premodified input would have worked as well (or even better) had it been as plentiful as the interactionally modified input.

A subsequent study (Pica, 1989) addressed this problem. This study compared the effects of interactionally generated input with premodified input which contained the same amount of original and repeated input and took the same length of time to present. In this case, then, the two conditions differed only in whether there were opportunities to negotiate. As Pica takes pains to point out the premodified input in the second study was based on the negotiated input rather than on the researcher's intuitions about what constituted comprehensible input. The premodified and interactionally adjusted input resulted in comprehension scores of 81% and 88% respectively - a difference that was not statistically significant. However, a post-hoc analysis of the results showed that for those learners who were rated as having lower comprehension ability by their teachers, the opportunity to interact was beneficial. Pica concludes that opportunities for negotiation may be most beneficial for learners in the early stages of L2 acquisition.'

What general conclusions can we reach regarding the second claim of the interaction hypothesis? At best, the research to date only suggests that simplified input helps comprehension. Also, much more work needs to be done before we can safely conclude if and when interactionally modified input works best. But, on balance, this part of the hypothesis looks promising. It is worth noting that L1 researchers (e.g. Ellis and Wells, 1981) have found that the input made available by parents with fast-learning children does not differ significantly on purely linguistic measures (i.e. what Long (1980) has called input features) from that made available by parents of slow-learning children, but differences are evident on a number of interactional measures (e.g. acknowledgements and expansions).

Perhaps the most convincing research has focussed on part three of the interaction hypothesis (i.e. the conditions that promote interactionally modified discourse). There is now ample testimony to the distorted nature of the discourse that occurs in teacher-dominated lessons and to the non-availablity of opportunities for modifying the structure of the classroom interactions that occur in them (cf. Long and Sato, 1983; Ellis, 1984). There have also been a number of interesting studies, building on Long's 1980 study, which have attempted to identify which general properties of tasks promote negotiation (e.g. Gass and Varonis, 1985; Pica and Doughty, 1985a and 1985b and various papers in Day, 1986). Again, though, it is wise not to place too much store on the results that have been obtained, as there are studies which also show that teacher-dominated discourse need not be totally devoid of interactional negotiation (cf. Van Lier, 1988), while much work remains to be done to establish which task factors are relevant to interaction and how these factors inter-relate in determining discourse outcomes. Also, research directed at part three of the hypothesis does not constitute a test of the essential elements of the hypothesis.

This review of the evidence that has been cited in support of the various parts of the interaction hypothesis indicates the need for caution. It is now a decade since the hypothesis was originally formulated and yet there is still no direct evidence to link interaction to acquisition and precious little to demonstrate that it promotes comprehension. This, perhaps, is not so surprising, nor very worrying, as SLA and, in particular, classroom SLA is till in its infancy. A decade, after all, is not long. Also, the absence of supportive research does not warrant the abandonment of a hypothesis that, in many ways, has contributed substantially to our current understanding of how learning takes place in the classroom context.

An Evaluation of the Interactional Hypothesis

We have now considered the nature of the theoretical claims advanced by the interactional hypothesis and have also examined some of the empirical evidence which has been cited in support of it. In this section, we will take a look at a number of theoretical objections and some counter-evidence to the first two parts' before arriving at a final evaluation of the hypothesis.

Is Comprehensible Input Necessary for Acquisition?

Many of the theoretical objections have focussed on part one of the hypothesis - the claim that comprehensible input is necessary for acquisition. Two rather different arguments, representing a weak and a strong attack on this position, have been advanced.

The weak objection is based on the view that although comprehensible input is necessary for acquisition it is not sufficient (which, as we have seen, is a point Long now acknowledges). Sharwood Smith (1986;241) argues that the processes of comprehension and acquisition are not the same. He suggests that input has a 'dual relevance' - there is input that will help the learner to interpret for meaning and there is input that she will use to advance her interlanguage. Sharwood Smith goes on to consider the different nature of the processes involved. In the case of comprehension, surface input is only briefly registered, as the learner rapidly recodes into 'deeper semantic and pragmatic codes' in which the message is then stored. In the case of acquisition the learner needs to undertake both a surface structure analysis and a semantic representation of the input. The input has to be held in memory sufficiently long for a comparison between its representation and whatever representation is provided by the rules of the learner's current grammar to be carried out. Without such a comparison no restructuring of the current grammar can take place. A somewhat similar position has been taken up by Faerch and Kasper (1986). They argue that acquisition only occurs when there is a 'gap' between the input and the learner's current knowledge and, crucially, when the learner perceives the gap as a gap in knowledge.

A characteristic of the weak objection is the rejection of any role for simplified input in L2 (or L1) acquisition. Sharwood Smith (1986) argues that simplified input functions as an aid to comprehension but not to acquisition. He refers to research by Bates (1982) which indicates that the characteristics of motherese (one kind of simplified input) are controlled by the child rather than the caretaker. White (1989) argues that simplified input will not help the learner to discover certain facts about the target language, such as the coreference possibilities of the English pronoun system. She points out that input consisting of simple sentences (e.g. 'Jane washed her') will not help the learner to discover how to treat pronouns in complex sentences (e.g. 'Jane watched television before she had dinner'). Both Sharwood Smith and White argue that simplified input is detrimental to language acquisition because it deprives the learner of useful structural information about the target language grammar.

These arguments are surely mistaken, particularly where simplified input derived from interactionally adjusted conversations is concerned.' Bates (1982), on whom Sharwood Smith sets considerable store, is wrong - there is evidence from Wells (1985) that parents increase the frequency of specific grammatical features just before these first appear in their children's speech. White is wrong to assume that simplified input is incapable of revealing facts such as the rules governing pronominal coreference. Consider the following constructed but totally feasible conversational exchange:

NS: Jane watched television before she had dinner last night.

NNS: Before who had dinner?

NS: Before Jane had dinner. Jane watched TV before Jane had dinner.

NNS: Oh, I see.

White seems to characterise simplified input entirely in terms of reduction in grammatical complexity and to ignore other aspects. But, as the above exchange illustrates, interactional modifications often work on quite complex strings, helping to make the grammatical relationships that exist transparent and, therefore, easier to acquire (cf. Pica, 1990, for examples of how this happens in real conversations involving learners). The simplified input derived from interactional modifications, then, certainly need not deprive the learner of useful structural data, as both Sharwood Smith and White claim.

It does not follow, of course, that modified input, even when interactionally derived, results in acquisition - as we have already noted in the previous section. Sato (1986) investigated the role of conversational modifications arising out of communication breakdown in two Vietnamese learner's naturalistic acquisition of past time reference (PTR) in English. Neither learner acquired past tense grammatical markers and Sato suggests that this might have been because both learners were able to communicate PTR effectively precisely because they were given substantial interactional support. In other words, the conversational interaction removed the need to acquire past tense grammatical markers. Sato concludes that conversational interaction may facilitate the acquisition of some structures but not others. This is an interesting conclusion because it conforms to the findings of other research which has shown that morphological aspects of English (such as past tense markers) are 'fragile', i.e. typically not acquired in

communicatively-rich learning environments) and, indeed, are resistant to acquisition even in advanced learners (cf. Bardovi-Harlig and Bofman, 1989).'

In rejecting Sharwood-Smith's and White's arguments regarding the null contribution of modified input, we need to be careful not to overstate its importance. At best, we can say that it may help acquisition. Also, the argument advanced by Sharwood Smith, Faerch and Kasper and Sato that the processes of comprehending and acquiring are distinct and that input/interaction that works for comprehension may not work for acquisition still stands. Apart from Sato's study already referred to, there is other evidence to support this position.

Ellis (forthcoming) asked two groups of upper intermediate adult learners (one Black South African and the other Indonesian) to read an 'English' text which had been rewritten so as to introduce a number of non-English grammatical rules (i.e. a zero plural morpheme rule and a number of German word order rules). The task was administered as a faster reading exercise and the learners' reading speeds and comprehension measured. On completion of the task the learners were asked to write down in as much detail as possible anything they had noticed about the grammar of the text they had read. A 'grammar score' for each learner was based on the accuracy and detail of their written comments. For the South Africans there was a statistically significant inverse relationship between reading speeds and comprehension scores on the one hand and grammar scores on the other. In other words, the faster readers and more successful comprehenders noticed less about the grammar. For, the Indonesians no such relationship was found. Ellis suggests that the difference between the two groups reflects differences in the way they approached the reading task. The Indonesians had received intensive training in how to read for meaning and, therefore, may have focussed primarily on comprehending the text. The South Africans had received no such training. Many read very slowly and may, therefore, have given themselves more opportunity to perceive the 'new' grammatical features.

Further evidence in support of the view that the availability of comprehensible input does not guarantee acquisition is provided by recent studies which have investigated grammar learning in 'communicative' classrooms (cf. Ellis, forthcoming for a review). Swain (1985) reported that early immersion students in Canada performed poorly on tests of grammatical competence. Hammerly (1987) reviewed six studies of French immersion, all of which showed that although the learners developed good listening and reading skills, they remained far from 'linguistically competent'. Hammerly argued that comprehensible input results

only in 'a very defective and probably terminal classroom pidgin' (p.397). Spada and Lightbown (1989) provide evidence to suggest that an intensive communicative ESL program in Quebec (5 hrs 5 days a week for 5 months) also resulted in low levels of grammatical acquisition (e.g. only 50% accuracy on plural-s, an early-acquired morpheme). These programs should not be considered failures, however, as they did result in considerable fluency and confidence in the learners' ability to use the L2. They suggest that comprehensible input is by itself not sufficient for acquisition of high levels of grammatical proficiency. One explanation for this is that the learners in such programs do not notice grammatical features in the input they comprehend. Such a conclusion is supported by the results of other studies (e.g. Spada, 1987) which suggest that drawing learners' conscious attention to grammatical form aids their acquisition, particularly if there is also access to comprehensible input through meaning-focussed communication.

Having considered the weak objection to the claim that comprehensible input is necessary for acquisition, we can now turn our attention to the strong objection. This disputes the very necessity of comprehensible input. This objection is based on learnability theory. White (1987), for instance, has argued that some grammatical features cannot be acquired purely on the basis of positive evidence (i.e. comprehensible input) and require negative input (i.e. feedback that draws conscious attention to the existence of the features). She notes that learners not only have to add new rules but also to lose transitional rules and that many of the latter cannot be disconfirmed purely on the basis of input, no matter how well-adjusted it is. Rutherford (1989) points out that learners manifest 'near universal failure to attain full target language competence' (p. 442) and that this may be because preemption is blocked in L2 acquisition. Rutherford, like white, argues it may be necessary to bring the difference between the learner's interim rules and the target language rules to the learner's conscious attention. Zobl (1983) and Eckman (1985) have both argued that acquisition entails 'generalisation'. That is, learners are credited with the ability to project grammatical knowledge they obtain from input to other, related grammatical knowledge not available or not attended to in the input . The gist of all these arguments is that input is not enough and that learners are equipped to go beyond the input.

The strong objection is able to muster some powerful arguments but little empirical evidence. In this respect it is not dissimilar to the interaction hypothesis itself. There is some evidence to support the capacity of learners to project beyond the input they receive (cf. Zobl, 1985; Eckman et al, 1988). But this evidence is open to challenge. The studies to date have not demonstrated that learners

generalize what they have learnt to new grammatical features, merely that they are able to increase the accuracy with which they perform features they have already learnt on the basis of input containing other features (cf. Jones, 1991, for a study that lends some support to this position). The studies have also only investigated a very limited number of grammatical features - relative clauses and possessive pronouns - between which distinct markedness relations are believed to exist. We do not know whether and to what extent similar relations hold between other structures. Also, to date, the studies that have investigated projection have involved inducing conscious attention to specific grammatical features through formal instruction rather than supplying access to these features through comprehensible input.

The strong and weak theoretical objections to the central claim of the interaction hypothesis and the empirical research which sustains them give rise to doubts as to (1) whether comprehensible input is sufficient for acquisition and (2) whether it is always necessary. In general, the case for (1) is stronger than that for (2). The input hypothesis, as currently formulated, is probably not tenable. However, as White (1987) has noted, some kind of input hypothesis is clearly necessary in any theory of L2 acquisition. Later a revised form of the input hypothesis will proposed.

Do Interactional Adjustments make Input Comprehensible?

The second claim of the interactional hypothesis - that modifications to the interactional structure of discourse promote comprehension - rests on firmer ground, although it too is in need of some revision.

We have already examined the study by Pica, Young and Doughty (1986) and Pica's (1989) follow up study, which lend some support to the claim that interactionally modified input aids comprehensions. A study by Loschky (1989) also indicates that opportunities for negotiating meaning are especially helpful where comprehension is concerned. Loschky investigated the acquisition of locative markers in L2 Japanese beginner learners. The learners were divided into three groups each of which received a different treatment. One group received baseline input,' another premodified input and a third interactionally modified input. Loschky tested the learner's comprehension of sentences containing the locative structures immediately after the treatment. He also tested their 'retention' of the structures by means of aural recognition and sentence verification tasks. The results

demonstrated a clear advantage for the group that received interactionally modified input where comprehension was concerned, but no advantage for retention. Loschky also found a non-significant correlation between the learners' gains in comprehension and vocabulary gains. This study, then, indicates that for beginner learners opportunities to negotiate meaning results in better comprehension, but has no apparent impact on acquisition. As such, it bears out the discussion in the previous section.

The claim that interactional modifications aid comprehension has met with opposition, however. First, there is some evidence to suggest that what appears to be the negotiation of meaning may not be anything of the kind. Hawkins (1985) carried out a study to determine whether the apparently appropriate responses made by learners in the course of negotiation actually signalled comprehension. She collected retrospective data from learners in order to determine whether their responses actually did represent comprehension and found that in 50% of the responses for which retrospective data were available comprehension had not in fact taken place.

Second, it has been pointed out that interactional modifications occur for other purposes than for negotiating meaning. Aston (1986) has argued convincingly for a social perspective on what he calls 'trouble shooting procedures' to complement the psycholinguistic perspective afforded by the interactional hypothesis. He notes that:

> Trouble shooting procedures can be employed to locate and deal with both troubles of accessibility and acceptability, and moreover can be used when trouble is neither present nor imminent.

Thus, modification to the structure of interaction occurs when the participants need to achieve 'a formal display of convergence'. In such cases, Aston suggests, they may go through a 'ritual of understanding or agreement' in order to show that the interaction has been successful. Aston goes on to argue that excessive trouble-shooting procedures may jeopardize communication from a social point of view.

Third, following on from Hawkins' and Aston's observations, the claim that it is the quantity of interactional modifications that matters has been challenged. Ehrlich, Avery and Yorio (1989;399) argue that 'the mere number of meaning negotiations within an interaction may not be a good predictor of the quality of comprehensible input'. In an interesting study, they compared the interactions of

eight native speaker/native speaker dyads with those of eight native speaker/non-native speaker dyads in a problem solving task requiring objects to be described. They found evidence for two relatively distinct styles in the speakers supplying the information. Skeletonizers provided the barest of details. Embroiderers tended to expand and embellish. Also, the skeletonizers tended to abandon negotiation when the descriptions strayed into a level of detail which the learners clearly could not handle, whereas the embroiderers carried on regardless. In the case of the NS/NS pairs, the style of the speaker did not affect overall success in the task, but in the case of the NS/NNS pairs it did; the skeletonizing pairs were more successful. Ehrlich, Avery and Yorio conclude that skeletonizing is more likely to result in comprehensible input. This study, then, indicates that it is not so much the quantity as the quality of interaction that counts. It goes some way to showing what particular aspects of negotiation facilitate comprehension.

Finally, we should also note that the research which has investigated interactional investigations has been extremely narrow. It has focussed on a very limited set of discourse functions found in negotiated interaction (cf. Table 1) and has neglected other aspects of discourse (such as topicalization) that may be important for comprehension. Also, it has been concerned primarily with discourse in English. We have no guarantee that the negotiation of meaning proceeds similarly in other languages. We do not know whether the same kinds of modifications occur cross-linguistically. A study by Kitazawa (1990) has reported considerable problems in identifying and classifying the interactional moves described by Long and Pica in Japanese conversations involving learners.

Two general conclusions seem possible regarding the relationship between interactional modifications and comprehension:

(1) Interactional modifications help learners to comprehend difficult material. They appear to aid comprehension to a greater extent than premodified input.

(2) Precisely when and how interactional modifications work for comprehension is still poorly understood, but it is becoming clear that it is the quality rather than the quantity that matters.

There is an obvious need for more studies (both descriptive and experimental) that probe the variables which determine whether and to what extent adjustments to discourse contribute to comprehension. Such studies will hopefully go beyond the

fairly restricted set of interactional features which have figured in the research to date.' They will also need to investigate the nature and the effects of modifications in languages other than English. In the meantime some revision of part two of the interactional hypothesis is in order.

Revising the Hypothesis

We have seen that only very limited evidence can currently be mustered in support of the interaction hypothesis. We have also seen that the hypothesis is open to serious challenge in two major ways. First, comprehensible input appears to be neither necessary nor sufficient for acquisition and second there appear to be conditions governing whether and when interactional modifications make input comprehensible. These constitute serious challenges, which cannot be ignored. One way forward might be to abandon the hypothesis and look for an alternative theory to explain how learners use input to build their interlanguages. This, in effect, is what White (1987) and others who view Universal Grammar as the key to explaining L2 acquisition have advocated. However, I do not think that abandonment is the best course or indeed necessary. For a start, the interactional hypothesis has made a valuable contribution to classroom second language research, motivating a number of studies that have addressed both the nature of classroom interaction and also its relationship to learners' understanding and learning. Also, the hypothesis has strong face validity. As many teachers have pointed out to me, it is common sense to work on the assumption that making learners understand what you say to them will help them learn the L2. Revision rather than abandonment of the hypothesis is the path I favour.

What is missing from the hypothesis is a theoretical account of <u>how</u> input made comprehensible through interactional modification results in acquisition. The revisions that I wish to propose are based on such an account.

The process of acquiring an L2 involves three basic procedures: (1) noticing, (2) comparison and (3) integration. Noticing entails the learner attending consciously to linguistic features in the input. Noticing entails perception and storage in short term memory. Thus, a feature that is noticed becomes 'preliminary intake' (Chaudron, 1985). Comparison also entails only short term memory. It involves the learner in identifying the difference between features noticed in the input and features currently in her own output. Integration takes place when the learner constructs new hypotheses on the basis of comparing input and her own

output and stores these in long term memory. Integration, therefore, results in 'final intake' (Chaudron, 1985). For acquisition to take place, all three procedures must occur. It does not follow that because a learner has noticed some feature in the input and has carried out a comparison with her own output that integration will take place. It is now generally recognised that psycholinguistic constraints of various kinds govern when new linguistic material is incorporated into the learner's interlanguage (cf. the Multidimensional Model, Pienemann, 1989). These constraints, working at the level of integration, are responsible for the developmental orders which have been observed in L2 acquisition. The position I seek to advance is that the role of input derived through interaction is primarily that of facilitating the processes of noticing and comparison.

According to Schmidt (1990), the process of noticing is necessarily a conscious one. He claims that 'you can't learn a foreign language through sublimal perception'. Conscious noticing, however, is not the same as deliberate attending. Noticing can take place either intentionally or inadvertently. Also, noticing does not require focal attention. When taking part in a conversation, for instance, the learner may be primarily focussed on the message content, but may also pay peripheral attention to striking linguistic features in the input. Under certain conditions, the learner may bring these features into focal attention. Lennon (1989) has documented how the advanced learners he studied appeared to switch their attention backwards and forwards from trying to communicate to trying to learn by consciously attending to input features even within the course of a single interaction. Schmidt argues that learners who notice the most, will learn the most.

The process of comparison is also potentially a conscious one, although it may often take place subconsciously. Schmidt and Frota (1986) carried out a detailed study of Schmidt's acquisition of Portuguese based on his entries in an admirably detailed diary. They illustrate how Schmidt often 'noticed the gap' between what he typically said and what he observed in the input. Here is an example of the kind of comparison Schmidt reported carrying out:

> I often say *dois anos antes* for 'two years ago'. I think it should be *anos atras.* I have been hearing it that way in conversation, I think.
>
> (Later the same day) I asked M which is correct and he says both are OK, but I am suspicious. Check with S tomorrow.

Schmidt and Frota provide several examples of how features that were noticed in the input and compared to current output subsequently appeared productively in Schmidt's conversational output. In fact, out of 21 verbal constructions they investigated, 20 were reported as having being noticed prior to their use in spontaneous production.

Of key interest is what induces a learner to notice features in the input and then to compare them to her current output. It is here that a role for comprehensible input and interaction appears likely. Modified input may be effective in drawing a learner's attention to features that would otherwise be ignored. This might occur in two principal ways - by increasing the frequency of specific forms at particular times (cf. Wells, 1985) and also by constructing messages in such a way that certain features become prominent in the input (e.g. by placing them in utterance initial position). Simplified input will only instigate noticing providing the learner is attending to it and this in turn is more likely to occur if the learner is able to understand the message content. As several studies have shown, beginners are not general successful in learning from unsimplified input such as that made available through TV or radio (cf. Snow et al, 1976). This is because such input is not comprehensible and so does not facilitate the noticing process. Simplified input does not ensure that noticing will take place, however. For one thing, if the input is simplified too much, there may be nothing new to notice. Also, the learner may be so focussed on message that she has no time to attend to linguistic features.

Negotiated interaction may be particularly useful in both helping the learner to notice new features and also to carry out a comparison with her existing output. Consider the following (again contrived) interaction:

NNS: No go disco this Saturday.

NS: Oh, so you're not going to the disco this Saturday?

NNS: Yeah, not going.

In such an exchange the native speaker reformulates the learner's utterance in the guise of a confirmation check and one feature of this reformulation is taken up in the learner's response. Such interactions afford the learner <u>overt</u> comparisons between interlanguage and target language forms. Of course, there is no guarantee that overt comparisons lead to the kind of <u>mental</u> comparison which Schmidt hypothesizes is necessary, but they surely create the conditions under which such

mental comparisons are more likely to occur. Thus modified interaction does not guarantee that noticing or comparison will take place; it merely facilitates it. Also, as we have already seen, certain kinds of modified interaction may work better for comprehension and thereby for noticing than others (cf. Ehrlich, Avery and Yorio, 1989).

The processes of noticing and comparison, then, may be facilitated when input is comprehensible and when interactional modification is possible. It is important to recognize, however, that there are likely to be a number of factors in addition to interactional modification that govern if and when these acquisitional processes take place (cf. Schmidt, 1990). The following is a provisional list of such factors:

(1) task demands (i.e. the instructional task causes the learner to attend to certain linguistic features because these are important for acquisition).

(2) unusual features (i.e. features that surprise the learner. Such features may work in similar ways to deviations in literary texts).

(3) markedness (i.e. features that are relatively unmarked may be easier to notice than features that are more marked).

(4) the learner's L1 (i.e. L2 features that match L1 features may be more noticeable, at least in the earlier stages of acquisition).

(5) individual learner differences (i.e. factors such as aptitude and motivation may influence whether or not a learner attends to new linguistic features in the input).

Whether or not these factors are important for noticing and comparing and in what ways remains an empirical question. What is important to recognise is that within the overall theory being proposed the availability of comprehensible input and of opportunities for negotiating meaning do not ensure that these processes will take place.

The role of comprehensible input and interactional modifications where integration of new linguistic material into interlanguage is concerned is less clear. The comprehensibility of input may prove to be irrelevant for integration. However, it is possible that interaction that enables the learner to use those features that have been noticed and compared in output may help. Such a position is

compatible with the more recent work on negotiated interaction which has sought to show how learners can be 'stretched' by obliging them to make their own output more comprehensible (see earlier discussion). But there is a strong likelihood that the integration of new structures depends to a large extent on factors other than input or output (e.g. innate knowledge of the universal properties of language and the learner's L1).

The theoretical position which has now been outlined suggests a number of revisions of the interactional hypothesis are in order. First, it is necessary to make a much weaker claim on behalf of comprehensible input. It must be seen as playing a facilitative rather than a necessary role in acquisition. Also due recognition must be given to the fact that the acquisition of some linguistic structures can occur independently of input. Second, modified input is seen as important for acquisition but only in the sense that it makes acquisition possible, not in the sense that it causes acquisition to take place. The conditions under which modified input may work need to be specified. Thus, it is hypothesized that modified input plays a part in enabling learners to carry out the preliminary processes of acquisition - noticing and comparison. The special usefulness of modified input derived from negotiating communication problems is also acknowledged. Third, a role for output is incorporated into the hypothesis, along the lines currently proposed by Pica and Long. Output is seen as a mechanism that facilitates the integration of new linguistic knowledge. It follows that situational conditions and tasks that promote interaction which produces comprehension and which encourages the processes of noticing, comparison and integration will be effective for acquisition. However, as little is currently known about what situations and tasks achieve this, no reference to external factors is incorporated into the revised hypothesis.

Conclusion

The focus of this paper has been the role played by face-to-face interaction in L2 acquisition. The interaction hypothesis, as proposed initially by Long and investigated by Pica, has served as the main theoretical account of this role. A number of problems with this hypothesis have been identified and a revised interaction hypothesis has been put forward. This claims that the comprehensible input derived from modified interaction may facilitate theprocesses of noticing linguistic features and comparing them with the features the learner derives from her current interlanguage. It also claims that learner production resulting from the attempt to negotiate meaning can facilitate the process of integrating new features

into interlanguage. These claims are weaker seen as neither necessary nor sufficient for acquisition. However, they are compatible with the results of empirical and theoretical enquiry into L2 acquisition.

---

(1) Comprehensible input facilitates L2 acquisition but is neither necessary not sufficient.

(2) Modifications to input, especially those which take place in the process of negotiating a communication problem make acquisition possible providing that the learners:

a. comprehend the input

b. notice new features in it and compare what is noticed with their own output.

(3) Interaction that requires learners to modify their initial output facilitates the process of integration.

Table 4: A revised version of the interactional hypothesis

One advantage of the revised hypothesis is that it is possible to see how it can be tested empirically. The absence of direct evidence in support of the original hypothesis is a reflection of the difficulty researchers have faced in testing it. Long (1985) suggested that the hypothesis might be tested indirectly by (1) showing that linguistic and conversational adjustments promote comprehension of input, (2) showing that comprehensible input promotes acquisition and (3) inferring that linguistic/conversational adjustments promote acquisition. But this suggestion has not proved very helpful as it is not clear how (2) can be achieved. The revised hypothesis does not face this difficulty because the relationship between input, comprehension and acquisition has now been more clearly defined. The nature of the link between modified input and the processes of noticing and comparing can be tested through introspection and retrospection because both of these processes take

place at a conscious level. One way is by asking learners to record what it is that they remember having noticed while they were performing a particular task and then examining the nature of the interactions in which the noticed features occurred (cf. Slimani, 1987, for an account such a methodology)."

Finally, I would like to join the many other researchers who have advised caution in applying the results of SLA studies to language pedagogy (cf. Lightbown. 1985). It is highly desirable, as Long (1987) has pointed out, that language pedagogy should be informed by research that is 'theoretically motivated' and has 'high valency' (i.e. capable of generalization beyond the specific context of the research site). The problem is that in such a young field as SLA there is no agreement regarding what constitutes a valid theory. As a result the research that has taken place, including the research based on the interaction hypothesis, affords few certainties. Such research cannot in my view be used to inform pedagogy. It is surely unwise, for instance, to propose that teachers select tasks according to how many interactional modifications they give rise to (cf. Long and Crookes, 1987), given how little we know about the relationship between these modifications and acquisition. The research should serve as a means (and not the only means) of illuminating language pedagogy by drawing attention to possible lines of intervention which teachers can then test out, ultimately accepting or dismissing them in the light of their classroom experience.

Notes:

1. Parker and Chaudron (1987) suggest that some characteristics of the speech addressed to learners do not reflect either input or interactional modification. They propose a third category of features which they label 'elaboration'. This includes modifications that involve redundancy (e.g. repetition of constituents, use of synonyms and rhetorical framing) and those that involve thematic structure (e.g. extraposition and cleft constructions).

2. Swain (1985) argues that output encourages acquisition by forcing learners to be precise, coherent and appropriate, by obliging them to process syntactically (as opposed to semantically) and by providing opportunities for them to revise hypotheses. The evidence cited in support of the output hypothesis is circumstantial (i.e. there is no direct evidence linking 'pushed output' to acquisition).

3. The Parker and chaudron (1987) study did attempt to establish to what extent modifications involving redundancy and thematic structure differentially contributed to reading comprehension, but found no effect for either.

4. Pica's (1989) study also investigated whether interactional modifications work as well for learners who listen to others interact as it does for those who actively participate. She found no statistical difference between the comprehension of the listeners and the participators.

5. The evaluation focuses on the first two parts of the hypothesis as it is these that are essential. Part three functions only as a corollary.

6. It is interesting to note that neither White nor Sharwood Smith make any mention of other forms of adjustment in input to learners (i.e. interactional or elaboration). They discuss only simplified input. It is not possible, of course, to dismiss a role for modified input by arguments that address only simplified input.

7. Vertical construction and incorporation (Hatch and Wagner-Gough, 1976) may serve as discourse strategies which help the learner to acquire syntactical structures. However, it is not clear how such strategies could facilitate the acquisition of morphological structures.

8. The baseline input in Loschky's study consisted of unmodified input derived from performing the task with native speakers.

9. Bygate's (1988) analysis of the kinds of interactions that take place in small group work involving learners of mixed proficiency levels suggests how the study of interaction can be taken further than the small set of discourse functions which the interaction hypothesis centres around. Bygate examined the forms and functions of what he calls 'satellite units' (i.e. chunks of language that are dependent on some previous part of the discourse). Bygate claims that these units facilitate the process of language acquisition, although he offers no evidence to demonstrate this.

10. Slimani recorded six lessons involving a non-native speaker teacher and intermediate learners of English in an Algerian technical college. She asked the learners to fill in an uptake chart, in which they recorded which items they had learnt during the lessons. Slimani then sought to identify where the items recorded occurred in each lesson and to examine the interactional properties that might have led to their recall. Slimani found no relationship between modified interaction and recall. However, given that the lessons were all very teacher-centred and afforded little opportunity for negotiation her study cannot be considered a good test of the revised interactional hypothesis.

REFERENCES

ALLWRIGHT, R., 1984: The Importance of Interaction in Classroom Language Learning, Applied Linguistics, 5, 156-71.

ASTON, G., 1986: Trouble Shooting in Interaction with Learners: The More the Merrier, Applied Linguistics, 7, 128-43.

BARDOVI-HARLIG, K., and Bofman, T., 1989: Attainment of Syntactic and Morphological Accuracy by Advanced Language Learners, Studies in Second Language Acquisition, 11, 17-34.

BYGATE, M., 1988: Units of Oral Expression and Language Learning in Small Group Interactions, Applied Linguistics, 9, 59-82.

CHAUDRON, C., 1983: Foreigner Talk in the Classroom. In Seliger, H and Long, M, (eds.) Classroom Oriented Research in Second Language Acquisition, Rowley, Mass., Newbury House.

CHAUDRON, C., 1985a: Comprehension, Comprehensibility and Learning in a Second Language, Studies in Second Language Acquisition, 7, 216-32.

CHAUDRON, C., 1985b: A Method for Examining the Input/intake Distinction. In S. Gass and C. Madden, (eds.)

CHAUDRON, C., 1988: Second Language Classrooms, Cambridge, Cambridge University Press.

DAHL, D., 1981: The Role of Experience in Speech Modifications for Second Language Learners, Minnesota Papers in Linguistics and Philosophy of Language, 7, 78-93.

DAY, R., ed. 1986: Talking to Learn: Conversation in Second Language Acquisition. Rowley, Mass. Newbury House.

ECKMAN, F., 1985: Some Theoretical and Pedagogical Implications of the Markedness Differential Hypothesis, Studies in Second Language Acquisition, 7, 289-307.

ECKMAN, F., Bell, L., and Nelson, D., 1988: On the Generalization of Relative Clause Instruction in the Acquisition of English as a Second Language, Applied Linguistics, 9, 1-20.

ERLICH, S., Avery, P., and Yorio, C., 1989: Discourse Structure and the Negotiation of Comprehensible Input, Studies in Second Language Acquisition, 11, 397-414.

ELLIS, R., 1984: Classroom Second Language Development, Oxford, Pergamon.

ELLIS, R., (forthcoming), Comprehension and the Acquisition of Grammatical Knowledge in a Second Language.

ELLIS, R., and Wells, G., 1980: Enabling Factors in Adult-child Discourse, First Language, 1, 46-82.

FAERCH, C., and Kasper, G., 1986: The Role of Comprehension in Second Language Acquisition, Applied Linguistics, 7, 257-74.

GASS, S., and Madden, C., (eds.), 1985: Input in Second Language Acquisition, Rowley, Mass. Newbury House.

GASS, S., and Varonis, E., 1985: Task Variation and Nonnative/nonnative Negotiation of Meaning. In Gass, S., and Madden, C., (eds.)

GLEITMAN, L., Newport, E., and Gleitman, H., 1984: The Current State of The Motherese Hypothesis, Journal of Child Language, 11:43-79.

HAMMERLY, H., 1987: The Immersion Approach: Litmus test of Second Language Acquisition through Classroom Communication, Modern Language Journal, 71, 395-401.

HATCH, E., 1978: Discourse Analysis and Second Language Acquisition. In E. Hatch (ed.), Second Language Acquisition, Rowley, Mass. Newbury House.

HATCH, E., and Wagner-Gough, J., 1976: Explaining Sequences and Variation in Second Language Acquisition, Language Learning, Special Issue 4, 39-57.

HAWKINS, B., 1985: Is an 'Appropriate Response' always so Appropriate? In S. Gass and C. Madden, (eds.)

KELCH, K., 1985: Modified Input as An Aid to Comprehension, Studies in Second Language Acquisition, 7, 81-90.

JOHNSON, P., 1981: Effects on Reading Comprehension of Language Complexity and Cultural Background of a Text, TESOL Quarterly, 15, 169-81.

JONES, A., 1991: Learning to Run before you Walk: Is it Feasible? Unpublished MEd paper, Tokyo, Temple University Japan.

KITAZAWA, M., 1990: Is Pica and Doughty's System of Interaction Classification Applicable to a Japanese Language Classroom? Unpublished DEd paper, Tokyo, Temple University Japan.

KRASHEN, S., 1977: Some Issues relating to the Monitor Model. In H. Brown, C. Yorio and R. Crymes, (eds.), On TESOL '77, 144-158, Washington, D.C., TESOL.

KRASHEN, S., 1980: The Theoretical and Practical Relevance of Simple Codes. In T. Scarcella and S. Krashen, (eds.), Research in Second Language Acquisition, Rowley, Mass., Newbury House.

KRASHEN, S., 1985: The Input Hypothesis. London: Longman.

LARSEN-FREEMAN, D., 1986: An Explanation for the Morpheme Accuracy Order of Learners of English of a Second Language, Language Learning, 26, 125-35.

LENNON, P., 1989: Introspection and Intentionality in Advanced Second-Language Acquisition, Language Learning, 39, 375-95.

LIGHTBOWN, P., 1983: Exploring Relationships between Developmental and Instructional Sequences. In H. Seliger and M. Long (eds.), Classroom Oriented Research in Second Language Acquisition, Rowley, Mass., Newbury House.

LIGHTBOWN, P., 1985: Great Expectations: Second-Language Acquisition Research and Classroom Teaching, Applied Linguistics, 6, 173-89.

LONG, M., 1980: Input, Interaction and Second Language Acquisition. Unpublished Ph.D. Dissertation, Los Angeles, UCLA.

LONG, M., 1983: Native Speaker/non-native Speaker Conversation in the Second Language Classroom. On TESOL '82: Pacific Perspectives on Language and Teaching. Eds. M. Clarke and J. Handscombe. Washington DC. : TESOL.

LONG, M. 1985: Input and Second Language Acquisition Theory. In Gass, S., and Madden, C., (eds.)

LONG, M., 1987: The Experimental Classroom, Annals, AAPSS, 490, 97-109.

*LONG, M., 1989: **Task, Group and Task-group Interactions**, University of Hawaii Working Papers in ESL, 8, 1-26.*

*LONG, M., and Crookes, G., 1987: **Intervention Points in Second Language Classroom Processes.** In B. Das, (ed.), Patterns in Classroom Interaction in Southeast Asia, Singapore University Press/RELC.*

*LONG, M., and Porter, R., 1985: **Group Work, Interlanguage Talk and Second Language Acquisition**, TESOL Quarterly, 19, 207-28.*

*LONG, M., and C. Sato. 1983: **Classroom Foreigner Talk Discourse: Forms and Functions of Teachers' Questions.** In H. Seliger and M. Long (eds.), Classroom Oriented Research in Second Language Acquisition, Rowley, Mass.: Newbury House.*

*LOSCHKY, L., 1989: **The Effects of Negotiated Interaction and Premodified Input on Second Language Comprehension and Retention.** Unpublished MA Thesis: University of Hawaii at Manoa.*

*PARKER, K., and Chaudron, C., 1987: **The Effects of Linguistic Simplification and Elaborative Modifications on L2 Comprehension**, University of Hawaii Working Papers in ESL, 6, 107-33.*

*PICA, T., 1987: **Second Language Acquisition, Social Interaction and the Classroom**, Applied Linguistics, 8, 1-25.*

*PICA, T., 1989: **Classroom Interaction, Participation and Comprehension: Redefining Relationships**, Papers in Applied Linguistics, University of Alabama, 1, 1-36.*

*PICA, T., 1990: **The Textual Outcomes of NS-NNS Negotiation: What do they Reveal about Second Language Learning?** Paper given at Conference on Text and Context: Cross-Disciplinary and Cross-Cultural Perspectives on Language Study, Cornell University, Ithaca, New York.*

*PICA, T., and Doughty, C., 1985a: **The Role of Group Work in Classroom Second Language Acquisition**, Studies in Second Language Acquisition, 7, 223-49.*

*PICA, T., and Doughty, C., 1985a: **Input and Interaction in the Communicative Classroom: A Comparison of Teacher-fronted and Group Activities.** In S. Gass and C. Madden (eds.)*

*PICA, T., R. Young and C. Doughty. 1987: **The Impact of Interaction on Comprehension.** TESOL Quarterly 21: 737-58.*

PIENEMANN, M., 1985: Learnability and Syllabus Construction. In K. Hyltenstam and M. Pienemann, (eds.), Modelling and Assessing Second Language Acquisition, Clevedon, Avon, Multilingual Matters.

PIENEMANN, M., 1989: Is Language Teachable? Psycholinguistic Experiments and Hypotheses, Applied Linguistics, 10, 52-79.

RUTHERFORD, W., 1989: Preemption and the Learning of L2 Grammars, Studies in Second Language Acquisition, 11, 441-458.

SATO, C., 1986: Conversation and Interlanguage Development: Rethinking the Connection. Ed. R. Day.

SCHMIDT, R., and Frota, S., 1986: Developing Basic Conversational Ability in a Second Language: A Case Study of an Adult Learner of Portuguese. In R. Day, (ed.)

SCHMIDT, R., 1990: The Role of Consciousness in Second Language Learning. Applied Linguistics 11,2.

SHARWOOD-SMITH, M., 1986: Comprehension vs. Acquisition: Two ways of Processing Input. Applied Linguistics 7: 239-56.

SLIMANI, A., 1989: The Role of Topicalization in Classroom Language Learning. System 17: 223-34.

SNOW, C., Arlman-Rupp, A., Hassing, Y., Jobse, J., Joosten, J., and Vorster, J., 1976: Mothers' Speech in Three Social Classes, Journal of Psycholinguistic Research, 5, 1-20.

SPADA, N., 1987: Relationships between Instructional Differences and Learning Outcomes: A Process-product Study of Communicative Language Teaching. Applied Linguistics 8: 137-55.

SPADA, N., and Lightbown, P., 1989: Intensive ESL Programmes in Quebec Primary Schools, TESL Canada Journal, 1, 11-32.

SPEIDEL, G., Tharp, R., and Kobayashi, L., 1975: Is there a Comprehension Problem for Children who speak Nonstandard English? A Study of Children with Hawaiian English backgrounds, Applied Psycholinguistics, 6, 83-96.

SWAIN, M., 1985: Communicative Competence: Some Roles of Comprehensible Input and Comprehensible Output in its Development. In H. Seliger and M. Long (eds.), Input in Second Language Acquisition, Rowley, Mass.: Newbury House.

VAN LIER, L., 1988: The Classroom and the Learner. London: Longman.

WELLS, G., 1985: Language Development in the Pre-school Years. Cambridge: Cambridge University Press.

WHITE, L., 1987: Against Comprehensible Input: The Input Hypothesis and the Development of Second Language Competence, Applied Linguistics, 8, 95-110.

WHITE, L., 1990: Second Language Acquisition and Universal Grammar, Studies in Second Language Acquisition, 12, 121-134.

ZOBL, H., 1983: Markedness and the Projection Problems, Language Learning, 32, 293-313.

ZOBL, H., 1985: Grammars in Search of Input and Intake. In S. Gass and C. Madden, (eds.)

LIST OF CONTRIBUTORS

Dr Kathleen M Bailey
Professor, TESOL MA Program and Director, Intensive English Program
Monterey Institute of International Studies
425 Van Buren Street
Monterey, CA 93940
U.S.A.

Dr Rod Bolitho
Deputy Director
International Education Centre
College of St Mark and St John
Derriford Road
Plymouth PL6 8BH
UNITED KINGDOM

Mr Martin Bygate
Centre for Applied Language Studies
Language Resource Centre
University of Reading
Whiteknights P O Box 218
Reading, Berkshire RG6 2AH
UNITED KINGDOM

Dr Andrew D Cohen
Dept of Linguistics
142 Klaeber Ct., U. of Minn.
320 16th Ave., S.E., Mpls., MN 55416
U.S.A.

Dr Rod Ellis
Professor of Applied Linguistics
Temple University Japan
1-16-7 Kami-Ochiai
Shinjuku-ku, Tokyo 161
JAPAN

Prof Stephen J Gaies
Visiting Professor
Temple University Japan
1-16-7 Kami-Ochiai
Shinjuku-ku, Tokyo 161
JAPAN

Dr Roger Griffiths
Associate Professor of English
Nagoya University of Commerce and Business Administration
Sagamine, Nisshin-cho
Aichi-gun, Aichi
JAPAN 470-01

Prof David Nunan
Associate Director
National Centre for English Language Teaching and Research
Macquarie University
Sydney, NSW 2109
AUSTRALIA

Dr N S Prabhu
Associate Professor
Dept of English Language and Literature
National University of Singapore
10 Kent Ridge Crescent
SINGAPORE 0511